

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

ATTORNEYS AT LAW

1100 NEW YORK AVENUE, N.W. • WASHINGTON, D.C. 20005-3934

PHONE (202) 371-2600 • FACSIMILE (202) 371-2540 • www.skgf.com

JOHN U.S. 11/28/00  
ROBERT GREENE STERNE  
WARD J. KESSLER  
EVA A. GOLDSTEIN  
SAMUEL L. FOX  
DAVID K.S. CORNWELL  
ROBERT W. ESMOND  
TRACY GENE S. DURKIN  
MICHELE A. CIMBALA  
MICHAEL B. RAY  
ROBERT E. SOKOHL  
ERIC K. STEFFE  
MICHAEL O. LEE

STEVEN R. LUDWIG  
JOHN M. COVERT\*\*  
LINDA E. ALCORN  
ROBERT C. MILLONIG  
MICHAEL V. MESSINGER  
JUDITH U. KIM  
TIMOTHY J. SHEA, JR.  
DONALD R. MCPHAIL  
PATRICK E. GARRETT  
STEPHEN S. WHITESIDE  
JEFFREY T. HELVEY\*\*  
HEIDI L. KRAUS

JEFFREY R. KURIN  
RAYMOND MILLIEN  
PATRICK D. O'BRIEN  
LAWRENCE B. BUGAISKY  
CRYSTAL D. SAYLES  
EDWARD W. YEE  
ALBERT L. FERRO\*  
DONALD R. BANOWIT  
PETER A. JACKMAN  
MOLLY A. MCCALL  
TERESA U. MEDLER  
JEFFREY S. WEAVER

KRISTIN K. VIDOVICH  
KENDRICK P. PATTERSON  
DONALD J. FEATHERSTONE  
GRANT E. REED  
VINCENT L. CAPUANO  
JOHN A. HARROUN\*  
ALBERT J. PASULO II\*  
W. BRIAN EDGE\*  
ELDORA ELLISON FLOYD\*  
W. RUSSELL SWINDELL  
THOMAS C. FIALA

KAREN R. MARKOWICZ\*\*  
SUZANNE E. ZISKA\*\*  
BRIAN J. DEL BUONO\*\*  
ANDREA J. KAMAE\*\*  
NANCY J. LEITH\*\*  
ELIZABETH J. HAANES\*\*  
MARK P. TERRY\*\*  
TARJA H. NAUKARINEN\*\*

\*BAR OTHER THAN D.C.  
\*\*REGISTERED PATENT AGENTS

November 28, 2000

WRITER'S DIRECT NUMBER:  
(202) 218-7825

INTERNET ADDRESS:  
SZISKA@SKGF.COM

Commissioner for Patents  
Washington, D.C. 20231

Box Patent Application

Re: U.S. Non-Provisional Utility Patent Application under 37 C.F.R. § 1.53(b)  
Appl. No. To Be Assigned; Filed: November 28, 2000  
For: **Increased Lysine Production by Gene Amplification**  
Inventors: Hanke *et al.*  
Our Ref: 1533.1030002

Sir:

The following documents are forwarded herewith for appropriate action by the U.S.  
Patent and Trademark Office:

1. USPTO Utility Patent Application Transmittal Form PTO/SB/05;
2. U.S. Utility Patent Application entitled:

**Increased Lysine Production by Gene Amplification**

and naming as inventors:

**Paul D. Hanke  
Lhing-Yew Li-D'Elia  
John Rayapati**

the application comprising:

Commissioner for Patents

November 28, 2000

Page 2

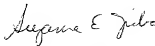
- a. specification containing:
  - i. 66 pages of description prior to the claims;
  - ii. 12 pages of claims ( 67 claims);
  - iii. a one (1) page abstract;
- b. 32 sheets of drawings: (Figures 1, 2, 3A, 3B, 4, 5A, 5B, 6-10, 11A, 11B, 12-14, 15A-15C, 16-20, 21A, 21B, 22-26);
3. 37 C.F.R. § 1.136(a)(3) Authorization to Treat a Reply As Incorporating An Extension of Time;
4. Fee Transmittal Letter PTO/SB/17;
5. Our check no. 29613 for \$1556.00 to cover:  
\$710.00 filing fee;  
\$846.00 claims over 20.
6. Two (2) return postcards.

It is respectfully requested that, of the two attached postcards, one be stamped with the filing date of these documents and returned to our courier, and the other, prepaid postcard, be stamped with the filing date and unofficial application number and returned as soon as possible.

This patent application is being submitted under 37 C.F.R. § 1.53(b) without Declaration and without filing fee.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.



Suzanne E. Ziska

Agent for Applicants

Registration No. 43,371

SEZ:vcf

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**UTILITY  
PATENT APPLICATION  
TRANSMITTAL**

(Only for new nonprovisional applications under 37 CFR § 1.53(b))

Attorney Docket No.	1533.1030002/SRL/SEZ
First Inventor	Paul D Hanke
Title	Increased Lysine Production by Gene Amplification
Express Mail Label No.	

 3541 US  
 09/12/04  
 11/28/00
**APPLICATION ELEMENTS**

See MPEP chapter 600 concerning utility patent application contents.

**ADDRESS TO:**
 Assistant Commissioner for Patents  
 Box Patent Application  
 Washington, DC 20231

1. ☒ Fee Transmittal Form (e.g., PTO/SB/17)  
(Submit an original, and a duplicate for fee processing)
2. ☐ Applicant claims small entity status.  
See 37 CFR 1.27.
3. ☒ Specification [Total Pages 79]  
(preferred arrangement set forth below)  
 - Descriptive title of the Invention  
 - Cross Reference to Related Applications  
 - Statement Regarding Derivative Sponsored R & D  
 - Reference to sequence listing, a table, or a computer program listing appendix  
 - Background of the Invention  
 - Brief Summary of the Invention  
 - Brief Description of the Drawings (if filed)  
 - Detailed Description  
 - Claim(s)  
 - Abstract of the Disclosure
4. ☒ Drawing(s) (35 U.S.C. 113) [Total Sheets 32]
5. Oath or Declaration [Total Pages     ]  
 a. ☐ Newly executed (original or copy)  
 b. ☐ Copy from a prior application (37 CFR 1.63(d))  
 (for continuation/divisional with Box 17 completed)  
 i. ☐ **DELETION OF INVENTOR(S)**  
 Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR §§ 1.63(d)(2) and 1.33(b).
6. ☐ Application Data Sheet. See 37 CFR 1.76.
7. ☐ CD-ROM or CD-R in duplicate, large table or Computer Program (Appendix)
8. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)  
 a. ☐ Computer Readable Form (CRF)  
 b. Specification Sequence Listing on:  
 i. ☐ CD-ROM or CD-R (2 copies), or  
 ii. ☐ paper  
 c. ☐ Statements verifying identity of above copies

**ACCOMPANYING APPLICATION PARTS**

9. ☐ Assignment Papers (cover sheet & document(s))
10. ☐ 37 CFR 3.73(b) Statement (when there is an assignee) ☐ Power of Attorney
11. ☐ English Translation Document (if applicable)
12. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
13. ☐ Preliminary Amendment
14. ☒ Two Return Receipt Postcard (MPEP 503)  
(Should be specifically itemized)
15. ☐ Certified Copy of Priority Document(s)  
(if foreign priority is claimed)
16. ☐ Request and Certification under 35 U.S.C. 122 (b)(2)(B)(i) Applicant must attach form PTO/SB/35 or its equivalent
17. ☒ Other: Authorization under 37 C.F.R. § 1.136(a)(3)

17. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment, or in an Application Data Sheet under 37 CFR 1.76:

☐ Continuation ☐ Divisional ☐ Continuation-in-Part (CIP) of prior application No:      /       
 Prior application information: Examiner      Group/Art Unit:     

For CONTINUATION OR DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 5b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only upon when a portion has been inadvertently omitted from the submitted application parts.

**18. CORRESPONDENCE ADDRESS**☐ Customer Number or Bar Code Label

(Insert Customer No. or Attach bar code label here)

or ☒ Correspondence address below

NAME	STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.				
ADDRESS	Attorneys at Law				
CITY	Washington	STATE	DC	ZIP CODE	20005-3934
COUNTRY	USA	TELEPHONE	(202) 371-2600	FAX	(202) 371-2540

NAME (Print/Type)	Suzanne E. Ziska	Registration No. (Attorney/Agent)	43,371
SIGNATURE	<i>Suzanne E. Ziska</i>	Date	November 28, 2000

Burden Hour Statement: this form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

# FEE TRANSMITTAL for FY 2001

Patent fees are subject to annual revision.

## Complete If Known

Application Number	To Be Assigned
Filing Date	November 28, 2000
First Named Inventor	Paul Hanke
Examiner Name	To Be Assigned
Group Art Unit	To Be Assigned
Attorney Docket No.	1533.1030002/SRL/SEZ

TOTAL AMOUNT OF PAYMENT (\$1556.00)

## METHOD OF PAYMENT (check one)

1. ☐ The Commissioner is hereby authorized to charge indicated fees and credit any overpayment to:

Deposit Account Number	19-0036
Deposit Account Name	Sterne, Kessler, Goldstein & Fox P.L.L.C.

- ☐ Charge Any Additional Fee Required  
Under 37 CFR §§ 1.16 and 1.17

- ☐ Applicant claims small entity status  
See 37 CFR 1.27

2. ☒ Payment Enclosed:

- ☒ Check ☐ Credit card ☐ Money Order ☒ Other\*  
\*Charge any deficiencies or credit any overpayments in the fees or fee calculations of Parts 1, 2 and 3 below to Deposit Account No. 19-0036

## FEE CALCULATION

### 1. BASIC FILING FEE

Large Fee Code (\$)	Entity Small Fee Code (\$)	Fee Description	Fee Paid
101	710 201 355	Utility filing fee	710.00
106	320 206 160	Design filing fee	
107	490 207 245	Plant filing fee	
108	710 208 355	Reissue filing fee	
114	150 214 75	Provisional filing fee	

SUBTOTAL (1) (\$) 710.00

### 2. EXTRA CLAIM FEES

Extra	Fee from below	Fee Paid
Total Claims 67 - 20** = 47	X 18	= 846.00
Indep. Claims 3 - 3** = 0	X	= 0
Multiple Dependent		=

Large Fee Code (\$)	Entity Small Fee Code (\$)	Fee Description
103	18 203 9	Claims in excess of 20
102	80 202 40	Independent claims in excess of 3
104	270 204 135	Multiple dependent claim
106	80 209 40	**Reissue independent claims over original patent
110	18 210 9	**Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) (\$) 846.00

\*\* or number previously paid, if greater. For Reissues, see above

## FEE CALCULATION (continued)

### 3. ADDITIONAL FEES

Large Fee Code (\$)	Entity Small Fee Code (\$)	Fee Description	Fee paid
105	130 205 65	Surcharge - late filing fee or oath	
127	50 227 25	Surcharge - late provisional filing fee or cover sheet	
139	136 139 130	Non-English specification	
147	2,320 147 2,320	For filing a request for ex parte reexamination	
112	920* 112 920*	Requesting publication of SIR prior to Examiner action	
113	1,840* 113 1,840*	Requesting publication of SIR after Examiner action	
115	110 215 55	Extension for reply within first month	
116	390 216 195	Extension for reply within second month	
117	890 217 445	Extension for reply within third month	
118	1,390 218 695	Extension for reply within fourth month	
128	1,890 228 945	Extension for reply within fifth month	
119	310 219 155	Notice of Appeal	
120	310 220 155	Filing a brief in support of an appeal	
121	270 221 135	Request for oral hearing	
138	1,510 138 1,510	Petition to institute a public use proceeding	
140	110 240 55	Petition to revive - unavoidable	
141	1,240 241 620	Petition to revive - unintentional	
142	1,240 242 620	Utility issue fee (or reissue)	
143	440 243 220	Design issue fee	
144	600 244 300	Plant issue fee	
122	130 122 130	Petitions to the Commissioner	
123	130 123 130	Petitions related to provisional applications	
126	180 126 180	Submission of Information Disclosure Stmt	
581	40 481 40	Recording each patent assignment per property (times number of properties)	
146	710 246 355	Filing a submission after final rejection (37 CFR 1.129(a))	
149	710 249 355	For each additional invention to be examined (37 CFR 1.129(b))	
179	710 279 355	Request for Continued Examination (RCE)	
169	900 169 900	Request for expedited examination of a design application	

Other fee (specify):

Other fee (specify):

\*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$) 0-

## SUBMITTED BY

Name (Print/Type)	Suzanne E. Ziska	Registration No. (Att/Reg/Agent)	43,371	Telephone	202-371-2600
Signature	<i>Suzanne E. Ziska</i>	Date	November 28, 2000		

## Complete (if applicable)

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2035.

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20531. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS.  
SEND TO: Assistant Commissioner for Patents, Washington, DC 20531.  
SGOF Rev. 11/13/00 msac

# Increased Lysine Production by Gene Amplification

Inventors: Hanke, Paul D.  
Li-D'Elia, Lhing-Yew  
Rayapati, John

5

## *Cross-Reference To Related Applications*

The present application claims benefit to the filing dates of U.S. Provisional Application No. 60/184,130, filed February 22, 2000; and U.S. Provisional Application No. 60/173,707, filed December 30, 1999, each of which is herein incorporated by reference.

10

## *Background of the Invention*

### *Field of the Invention*

15

The invention relates to the areas of microbial genetics and recombinant DNA technology. The invention provides gene sequences, vectors, microorganisms, promoters and regulatory proteins useful for the production of L-lysine. The invention further provides a method to increase the production of L-lysine.

### *Related Art*

20

L-lysine is an important economic product obtained principally by industrial-scale fermentation utilizing the Gram positive *Corynebacterium glutamicum*, *Brevibacterium flavum* and *Brevibacterium lactofermentum* (Kleemann, A., *et. al.*, Amino Acids, in ULLMANN'S ENCYCLOPEDIA OF INDUSTRIAL CHEMISTRY, vol. A2, pp. 57-97, Weinham: VCH-Verlagsgesellschaft (1985)).

25

The stereospecificity of the amino acids produced by fermentation makes the process advantageous compared with synthetic processes; generally L-form

amino acids are produced by the microbial fermentation process. The production of L-lysine and other amino acids through fermentation, utilizing cheap carbon sources such as molasses, glucose, acetic acid and ethanol, is a relatively inexpensive means of production.

Microorganisms employed in microbial processes for amino acid production may be divided into 4 classes: wild-type strain, auxotrophic mutant, regulatory mutant and auxotrophic regulatory mutant (K. Nakayama *et al.*, in NUTRITIONAL IMPROVEMENT OF FOOD AND FEED PROTEINS, M. Friedman, ed., (1978), pp. 649-661).

Several fermentation processes utilizing various strains isolated for auxotrophic or resistance properties are known in the art for the production of L-lysine: U.S. Patent No. 2,979,439 discloses mutants requiring amino acid supplementation (homoserine, or L-methionine and L- threonine); U.S. Patent No. 3,700,557 discloses mutants having a nutritional requirement for L- threonine, L-methionine, L-arginine, L-histidine, L-leucine, L-isoleucine, L- phenylalanine, L-cystine, or L-cysteine; U.S. Patent No. 3,707,441 discloses a mutant having a resistance to an L-lysine analog; U.S. Patent No. 3,687,810 discloses a mutant having both an ability to produce L-lysine and a resistance to bacitracin, penicillin G or polymyxin; U.S. Patent No. 3,708,395 discloses mutants having a nutritional requirement for homoserine, L-threonine, L- threonine and L-methionine, L-leucine, L-isoleucine or mixtures thereof and a resistance to L-lysine, L-threonine, L-isoleucine or analogs thereof; U.S. Patent No. 3,825,472 discloses a mutant having a resistance to an L-lysine analog; U.S. Patent No. 4,169,763 discloses mutant strains of *Corynebacterium* that produce L-lysine and are resistant to at least one of aspartic analogs and sulfa drugs; U.S. Patent No. 5,846,790 discloses a mutant strain able to produce L-glutamic acid and L-lysine in the absence of any biotin action-suppressing agent; and U.S. Patent No. 5,650,304 discloses a strain belonging to the genus *Corynebacterium* or *Brevibacterium* for the production of L-lysine that is resistant to 4-N-(D-alanyl)-2,4-diamino-2,4-dideoxy-L-arabinose 2,4-dideoxy-L-arabinose or a derivative thereof.

A considerable amount is known regarding the biochemical pathway for L-lysine synthesis in *Corynebacterium* species (recently reviewed by Sahn *et al.*, *Ann. N. Y. Acad. Sci.* 782: 25-39 (1996)). Entry into the L-lysine pathway begins with L-aspartate (see Figure 1), which itself is produced by transamination of oxaloacetate. A special feature of *C. glutamicum* is its ability to convert the L-lysine intermediate piperidine 2,6-dicarboxylate to diaminopimelate by two different routes, i.e. by reactions involving succinylated intermediates or by the single reaction of diaminopimelate dehydrogenase. Overall, carbon flux into the pathway is regulated at two points: first, through feedback inhibition of aspartate kinase by the levels of both L-threonine and L-lysine; and second through the control of the level of dihydrodipicolinate synthase. Therefore, increased production of L-lysine may be obtained in *Corynebacterium* species by deregulating and increasing the activity of these two enzymes.

More recent developments in the area of L-lysine fermentative production in *Corynebacterium* species involve the use of molecular biology techniques to augment L-lysine production. The following examples are provided as being exemplary of the art: U. S. Patent Nos. 4,560,654 and 5,236,831 disclose an L-lysine producing mutant strain obtained by transforming a host *Corynebacterium* or *Brevibacterium* species microorganism which is sensitive to S-(2-aminoethyl)-cysteine with a recombinant DNA molecule wherein a DNA fragment conferring both resistance to S-(2-aminoethyl)-cysteine and L-lysine producing ability is inserted into a vector DNA; U. S. Patent No. 5,766,925 discloses a mutant strain produced by integrating a gene coding for aspartokinase, originating from coryneform bacteria, with desensitized feedback inhibition by L-lysine and L-threonine, into chromosomal DNA of a *Corynebacterium* species bacterium harboring leaky type homoserine dehydrogenase or a *Corynebacterium* species deficient in homoserine dehydrogenase gene; increased L-lysine production is obtained by gene amplification by way of a plasmid vector or utilizing a gene replacement strategy. European Patent Applications EP 0 811 682 A2 and EP 0 854 189 A2

both provide for increased production of L-lysine in *Corynebacterium* species by way of gene amplification based on plasmid copy number.

### ***Summary of the Invention***

5 It is an object of the invention to provide a method to increase the production of an amino acid in *Corynebacterium* species by amplifying, i.e., increasing, the number of a gene or genes of an amino acid biosynthetic pathway in a host cell. Particularly preferred *Corynebacterium* species include *Corynebacterium glutamicum*, *Brevibacterium flavum*, and *Brevibacterium lactofermentum*.

10 It is an object of the invention to provide an isolated feed back resistant aspartokinase enzyme wherein the naturally occurring threonine amino acid residue 380 in the feedback sensitive form is changed to isoleucine in the *ask* gene of ATCC 21529. It is an object of the invention to provide an isolated ask polypeptide comprising the amino acid sequence of SEQ ID NO:2. It is another  
15 object of the invention to provide an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide sequence of SEQ ID NO:2. It is another object of the invention to provide an isolated polynucleotide molecule comprising a nucleic acid having the sequence of SEQ ID NO:1.

20 It is another object of the invention to provide a method comprising transforming a *Corynebacterium* species host cell with a polynucleotide molecule comprising a nucleotide sequence encoding a polypeptide comprising amino acid SEQ ID NO:2, wherein said isolated polynucleotide molecule is integrated into said host cell's chromosome thereby increasing the total number of said amino acid biosynthetic pathway genes in said host cell chromosome, and selecting a  
25 transformed host cell. It is a further object of the invention to provide a method comprising screening for increased amino acid production. The method may further comprise growing said transformed host cell in a medium and purifying an amino acid produced by said transformed host cell.



In another embodiment, a method to increase the production of an amino acid is a method comprising transforming a *Corynebacterium* species host cell with an isolated nucleic acid molecule encoding the amino acid sequence of SEQ ID NO:2, wherein said isolated nucleic acid molecule is integrated into said host cell's chromosome thereby increasing the total number of said amino acid biosynthetic pathway genes in said host cell chromosome, and wherein said isolated nucleic acid molecule further comprises at least one of the following: a polynucleotide encoding a *Corynebacterium* species lysine pathway *asd* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *dapA* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *dapB* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *ddh* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *lysA* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *lysA* amino acid sequence; a polynucleotide encoding a *Corynebacterium* species lysine pathway *ORF2* amino acid sequence, and selecting a transformed host cell. The method may further comprise growing said transformed host cell in a medium and purifying an amino acid produced by said transformed host cell.

The term " *lysA* " refers to a truncated *lysA* gene or amino acid sequence used by Applicants and described *infra*. The term "*lysA*" refers to the full length *lysA* gene or amino acid sequence used by Applicants and described *infra*.

It is another object of the invention to provide an isolated polynucleotide molecule comprising a nucleic acid molecule encoding the *Corynebacterium glutamicum* lysine pathway *ask* amino acid sequence of SEQ ID NO:2; and at least one additional *Corynebacterium* species lysine pathway gene selected from the group consisting of a nucleic acid molecule encoding the *asd* polypeptide, a nucleic acid molecule encoding the *dapA* polypeptide, a nucleic acid molecule encoding the *dapB* polypeptide, a nucleic acid molecule encoding the *ddh* polypeptide, a nucleic acid molecule encoding the *lysA* polypeptide, a nucleic acid molecule encoding the *lysA* polypeptide and a nucleic acid molecule

encoding the *ORF2* polypeptide. In a preferred embodiment of the invention, the isolated polynucleotide molecule comprises pK184-KDABH'L. In another preferred embodiment of the invention, the isolated nucleic acid molecule comprises pK184-KDAB. In another preferred embodiment of the invention, the isolated nucleic acid molecule comprises pD2-KDABHL. In another preferred embodiment of the invention, the isolated nucleic acid molecule comprises pD11-KDABH'L.

It is another object of the invention to provide a host cell transformed with an isolated polynucleotide molecule comprising a nucleotide sequence encoding an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the isolated nucleic acid molecule is integrated into the host cell's chromosome thereby increasing the total number of amino acid biosynthetic pathway genes in the host cell chromosome. In one embodiment the polynucleotide further comprises at least one additional *Corynebacterium* species lysine pathway gene selected from the group consisting of: a nucleic acid molecule encoding an *asd* polypeptide; a nucleic acid molecule encoding a *dapA* polypeptide; a nucleic acid molecule encoding a *dapB* polypeptide; a nucleic acid molecule encoding a *ddh* polypeptide; a nucleic acid molecule encoding a *'lysA* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; and a nucleic acid molecule encoding an *ORF2* polypeptide.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule encoding a polypeptide wherein said *asd* polypeptide is SEQ ID NO:4; said *dapA* polypeptide is SEQ ID NO:6; said *dapB* polypeptide is SEQ ID NO:8; said *ddh* polypeptide is SEQ ID NO:10; said *'lysA* polypeptide is SEQ ID NO: 21; said *lysA* polypeptide is SEQ ID NO:14; and said *ORF2* polypeptide is SEQ ID NO:16.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule wherein said *asd* polypeptide is SEQ ID NO:4; said *dapA* polypeptide is SEQ ID NO:6; said *dapB* polypeptide is SEQ ID NO:8; said *ddh* polypeptide is SEQ ID NO:10; said *'lysA* polypeptide is SEQ ID NO:21; said

*lysA* polypeptide is SEQ ID NO:14; and said *ORF2* polypeptide is SEQ ID NO:16.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another embodiment, the polynucleotide further comprises a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In one embodiment, the transformed host cell is a *Brevibacterium* selected from the group consisting of *Brevibacterium flavum* NRRL-B30218, *Brevibacterium flavum* NRRL-B30219, *Brevibacterium lactofermentum* NRRL-B30220, *Brevibacterium lactofermentum* NRRL-B30221, *Brevibacterium lactofermentum* NRRL-B30222, *Brevibacterium flavum* NRRL-30234 and *Brevibacterium lactofermentum* NRRL-30235. In another embodiment, the host cell is *Escherichia coli* DH5  $\alpha$  MCR NRRL-B30228. In another embodiment, the host cell is a *C. glutamicum* selected from the group consisting of *C. glutamicum* NRRL-B30236 and *C. glutamicum* NRRL-B30237.

It is another object of the invention to provide a method of producing lysine comprising culturing the host cells comprising the amino acid sequence of SEQ ID NO: 2 wherein said host cells comprise one or more of (a) increased enzyme activity of one or more lysine biosynthetic pathway enzymes compared to the genetically unaltered nonhuman host cell; (b) one or more copies of each gene encoding a lysine biosynthetic pathway enzyme; and, (c) alteration of one or more transcription factors regulating transcription of one or more genes encoding a lysine biosynthetic pathway enzyme, wherein said host cell produces lysine in said culture medium. In one embodiment of the invention, the increased enzyme activity comprises overexpressing one or more genes encoding one or more lysine biosynthetic pathway enzymes. In another embodiment of the invention the increased enzyme activity results from the activity of one or more modified lysine biosynthetic pathway enzymes wherein said enzyme modification results in a change in kinetic parameters, allosteric regulation, or both, compared to the enzyme lacking the modification. In another embodiment of the invention, alteration of one or more transcription factors comprises one or more mutations in transcription inhibitor proteins, one or more mutations in transcription activator proteins, or both, wherein said one or more mutations increases transcription of the target nucleotide sequence compared to the transcription by said one or more transcription factors lacking said alteration(s).

It is an object of the invention to provide an isolated polypeptide, wherein said polypeptide comprises an amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:19. It is a further object of the invention to provide an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:19. It is a further object of the invention to provide an isolated polynucleotide comprising a nucleic acid having the sequence of SEQ ID NO:18. It is another object of the invention to provide host cell NRRL B30360.

It is an object of the invention to provide an isolated polypeptide wherein said polypeptide comprises a polypeptide having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21. It is a further object of the invention to provide an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:21. It is a further object of the invention to provide a polynucleotide molecule comprising a nucleic acid having the sequence of SEQ ID NO:20.

It is an object of the invention to provide an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, further comprising a promoter sequence where said promoter sequence has at least 95% sequence identity to SEQ ID NO:17. It is a further object of the invention to provide an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polynucleotide molecule further comprises the sequence of SEQ ID NO:17. It is a further object of the invention to provide a host cell NRRL B30359.

Further objects and advantages of the present invention will be clear from the description that follows.

### ***Brief Description of the Figures***

**Figure 1.** A schematic of the L-lysine biosynthetic pathway in *Corynebacterium glutamicum* (Sahm *et al.*).

**Figure 2.** The nucleotide sequence of *ask* (ATCC 21529 sequence) (SEQ ID NO:1).

**Figure 3 A, B.** The amino acid sequence of *ask* (ATCC 21529 sequence) (SEQ ID NO:2).

**Figure 4.** The nucleotide sequence of *asd* (ATCC 21529 sequence) (SEQ ID NO:3).

**Figure 5 A, B.** The amino acid sequence of *asd* (ATCC 21529 sequence) (SEQ ID NO:4).

**Figure 6.** The nucleotide sequence of *dapA* (NRRL-B11474) (SEQ ID NO:5).

**Figure 7.** The amino acid sequence of *dapA* (NRRL-B11474) (SEQ ID NO:6).

**Figure 8.** The nucleotide sequence of *dapB* (NRRL-B11474) (SEQ ID NO:7).

**Figure 9.** The amino acid sequence of *dapB* (NRRL-B11474) (SEQ ID NO:8).

**Figure 10.** The nucleotide sequence of *ddh* (NRRL-B11474) (SEQ ID NO:9).

**Figure 11 A, B.** The amino acid sequence of *ddh* (NRRL-B11474) (SEQ ID NO:10).

**Figure 12.** The nucleotide sequence of full length *lysA* (NRRL-B11474) (SEQ ID NO:11) used to obtain the truncated *lysA* ('*lysA*') nucleotide sequence. Underlined region annealed with *lysA* primer.

**Figure 13.** The amino acid sequence of full length *lysA* (NRRL-B11474) (SEQ ID NO:12) comprising the truncated *lysA* ('*lysA*') amino acid sequence (SEQ ID NO: 21). Underlined L: the last amino acid residue of *lysA* encoded in the truncated PCR product.

**Figure 14.** The nucleotide sequence of full length *lysA* (pRS6) (SEQ ID NO:13).

**Figure 15 A, B, C.** The amino acid sequence of full length *lysA* (pRS6) (SEQ ID NO:14).

**Figure 16.** The nucleotide sequence of ORF2 (NRRL-B11474) (SEQ ID NO:15).

**Figure 17.** The amino acid sequence of ORF2 (NRRL-B11474) (SEQ ID NO:16).

**Figure 18.** A schematic depiction of the construction of the 5 and 6 lysine pathway gene constructs of the invention.

**Figure 19.** Comparison of the aspartokinase (*ask*) amino acid sequence from ATCC13032, N13 and ATCC21529.

**Figure 20.** The nucleotide sequence of the HpaI-PvuII fragment from pRS6 (SEQ ID NO:17) comprising the P1 promoter.

**Figure 21 A, B.** A schematic depiction of the construction of the pDElia2-KDABHP1L construct.

**Figure 22.** A schematic depiction of the construction of the pDElia2<sub>EC5</sub>-KDBHL construct.

**Figure 23.** The nucleotide sequence of truncated ORF2 (SEQ ID NO:18).

**Figure 24.** The amino acid sequence of truncated ORF2 (SEQ ID NO:19).

**Figure 25.** The nucleotide sequence of truncated LysA ('lysA')(NRRL-B11474) (SEQ ID NO:20).

**Figure 26.** The amino acid sequence of truncated LysA ('LysA')(NRRL-B11474) (SEQ ID NO:21).

## ***Detailed Description of the Preferred Embodiments***

### ***A. Definitions***

In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided. It is also to be noted that the term "a" or "an" entity, refers to one or more of that entity; for example, "a polynucleotide," is understood to represent one or more polynucleotides.

**Allosteric Regulation.** As used herein, the term refers to regulation of enzyme activity through the binding of one or more ligands (allosteric effectors) to one or more binding sites. The ligands may be the same molecule or different molecules. The molecules bind to sites on the enzyme other than the enzyme active site. As a result of the binding, a conformational change is induced in the enzyme which regulates affinity of the active site for its substrate or other ligands. Allosteric effectors may serve to enhance catalytic site substrate affinity (allosteric activators) or to reduce affinity (allosteric repressors). Allosteric effectors form the basis of metabolic control mechanisms such as feedback loops, for example (See, Copeland, Robert A., in *Enzymes. A Practical Introduction to Structure, Mechanism, and Data Analysis*, pages 279-296, Wiley-VCH, New York (1996)).

**Amino Acid Biosynthetic Pathway Genes.** As used herein, the term "amino acid biosynthetic pathway gene(s)" is meant to include those genes and genes fragments encoding peptides, polypeptides, proteins, and enzymes, which are directly involved in the synthesis of amino acids. These genes may be identical to those which naturally occur within a host cell and are involved in the synthesis of any amino acid, and particularly lysine, within that host cell. Alternatively, there may be modifications or mutations of such genes, for example, the genes may contain modifications or mutations which do not significantly affect the biological activity of the encoded protein. For example, the natural gene may be modified by mutagenesis or by introducing or



substituting one or more nucleotides or by removing nonessential regions of the gene. Such modifications are readily performed by standard techniques.

**Auxotroph.** As used herein, the term refers to a strain of microorganism requiring for growth an external source of a specific metabolite that cannot be synthesized because of an acquired genetic defect.

**Amino Acid Supplement.** As used herein, the term refers to an amino acid required for growth and added to minimal media to support auxotroph growth.

**Chromosomal Integration.** As used herein, the term refers to the insertion of an exogenous DNA fragment into the chromosome of a host organism; more particularly, the term is used to refer to homologous recombination between an exogenous DNA fragment and the appropriate region of the host cell chromosome.

**Enhancers.** As used herein, the term refers to a DNA sequence which can stimulate promoter activity and may be an endogenous element or a heterologous element inserted to enhance the level, i.e., strength of a promoter.

**High Yield Derivative.** As used herein, the term refers to strain of microorganism that produces a higher yield from dextrose of a specific amino acid when compared with the parental strain from which it is derived.

**Host Cell.** As used herein, the term "host cell" is intended to be interchangeable with the term "microorganism." Where a difference is intended, the difference will be made clear.

**Isolated Nucleic Acid Molecule.** As used herein, the term is intended to mean a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid

molecules according to the present invention further include such molecules produced synthetically.

**Lysine Biosynthetic Pathway Protein.** As used herein, the term "lysine biosynthetic pathway protein" is meant to include those peptides, polypeptides, proteins, and enzymes, which are directly involved in the synthesis of lysine from aspartate. Also included are amino acid sequences as encoded by open reading frames (ORF), where the ORF is associated with a lysine biosynthetic pathway operon. These proteins may be identical to those which naturally occur within a host cell and are involved in the synthesis of lysine within that host cell. Alternatively, there may be modifications or mutations of such proteins, for example, the proteins may contain modifications or mutations which do not significantly affect the biological activity of the protein. For example, the natural protein may be modified by mutagenesis or by introducing or substituting one or more amino acids, preferably by conservative amino acid substitution, or by removing nonessential regions of the protein. Such modifications are readily performed by standard techniques. Alternatively, lysine biosynthetic proteins may be heterologous to the particular host cell. Such proteins may be from any organism having genes encoding proteins having the same, or similar, biosynthetic roles.

**Mutagenesis.** As used herein, the term refers to a process whereby a mutation is generated in DNA. With "random" mutagenesis, the exact site of mutation is not predictable, occurring anywhere in the genome of the microorganism, and the mutation is brought about as a result of physical damage caused by agents such as radiation or chemical treatment. rDNA mutagenesis is directed to a cloned DNA of interest, and it may be random or site-directed.

**Mutation.** As used herein, the term refers to a one or more base pair change, insertion or deletion, or a combination thereof, in the nucleotide sequence of interest.

**Operably Linked.** As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. A nucleic acid is "operably linked" when it is placed into a functional relationship with another

nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous and, where necessary, join two protein coding regions, contiguous and in reading frame. However, since enhancers generally function when separated from the promoter by several kilobases and intronic sequences may be of variable lengths, some polynucleotide elements may be operably linked but not contiguous.

**Operon.** As used herein, the term refers to a contiguous portion of a transcriptional complex in which two or more open reading frames encoding polypeptides are transcribed as a multi-cistronic messenger RNA, controlled by a cis-acting promoter and other cis-acting sequences necessary for efficient transcription, as well as additional cis acting sequences important for efficient transcription and translation (*e.g.*, mRNA stability controlling regions and transcription termination regions). The term generally also refers to a unit of gene expression and regulation, including the structural genes and regulatory elements in DNA.

**Parental Strain.** As used herein, the term refers to a strain of host cell subjected to some form of treatment to yield the host cell of the invention.

**Percent Yield From Dextrose.** As used herein, the term refers to the yield of amino acid from dextrose defined by the formula  $[(\text{g amino acid produced} / \text{g dextrose consumed}) * 100] = \% \text{ Yield}$ .

**Phenotype.** As used herein, the term refers to observable physical characteristics dependent upon the genetic constitution of a host cell.

**Promoter.** As used herein, the term "promoter" has its art-recognized meaning, denoting a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription and thus refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes. In general, a coding sequence is located 3' to a promoter sequence. Sequence elements within promoters that function in the

initiation of transcription are often characterized by consensus nucleotide sequences. The promoter sequence consists of proximal and more distal upstream elements (enhancers). As used herein, the term "endogenous promoter" refers to a promoter sequence which is a naturally occurring promoter sequence in that host microorganism. The term "heterologous promoter" refers to a promoter sequence which is a non-naturally occurring promoter sequence in that host microorganism. The heterologous occurring promoter sequence may be from any prokaryotic or eukaryotic organism. A synthetic promoter is a nucleotide sequence, having promoter activity, and not found naturally occurring in nature.

Promoters may be derived in their entirety from a native gene, or be hybrid promoters. Hybrid promoters are composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. Hybrid promoters may be constitutive, inducible or environmentally responsive.

Useful promoters include constitutive and inducible promoters. Many such promoter sequences are known in the art. See, for example, U.S. Pat. Nos. 4,980,285; 5,631,150; 5,707,828; 5,759,828; 5,888,783; 5,919,670, and, *Sambrook, et al., Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Press (1989). Other useful promoters include promoters which are neither constitutive nor responsive to a specific (or known) inducer molecule. Such promoters may include those that respond to developmental cues (such as growth phase of the culture), or environmental cues (such as pH, osmoticum, heat, or cell density, for example).

Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, or the presence of light. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different cell types, or in response to different environmental conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters." It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined,

DNA fragments of different lengths may have identical or similar promoter activity.

**Relative Growth.** As used herein, the term refers to a measurement providing an assessment of growth by directly comparing growth of a parental strain with that of a progeny strain over a defined time period and with a defined medium.

**Transcription factor.** As used herein, the term "transcription factor" refers to RNA polymerases, and other proteins that interact with DNA in a sequence-specific manner and exert transcriptional regulatory effects. Transcriptional factors may be transcription inhibitory proteins or transcription activator proteins. In the context of the present invention, binding sites for transcription factors (or transcription complexes) are often included in the transcriptional regulatory element(s).

**Transcription factor recognition site.** As used herein, a "transcription factor recognition site" and a "transcription factor binding site" refer to a polynucleotide sequence(s) or sequence motif(s) which are identified as being sites for the sequence-specific interaction of one or more transcription factors, frequently taking the form of direct protein-DNA binding. Typically, transcription factor binding sites can be identified by DNA footprinting, gel mobility shift assays, and the like, and/or can be predicted on the basis of known consensus sequence motifs, or by other methods known to those of skill in the art.

**Transcriptional Complex.** As used herein, the term "transcriptional unit" or "transcriptional complex" refers to a polynucleotide sequence that comprises a structural gene (one or more exons), a cis-acting linked promoter and one or more other cis-acting sequences necessary for efficient transcription of the structural sequences, distal regulatory elements necessary for appropriate transcription of the structural sequences, and additional cis sequences important for efficient transcription and translation (e.g., polyadenylation site, mRNA stability controlling sequences). See, for example U.S. Patent No. 6,057,299.

**Transcriptional Regulatory Element.** As used herein, the term "transcriptional regulatory element" refers to a DNA sequence which activates transcription alone or in combination with one or more other DNA sequences. A transcriptional regulatory element can, for example, comprise a promoter, response element, negative regulatory element, silencer element, gene suppressor, and/or enhancer. See, for example, U.S. Patent No. 6,057,299.

**B. Microbiological and Recombinant DNA Methodologies**

The invention as provided herein utilizes some methods and techniques that are known to those skilled in the arts of microbiology and recombinant DNA technologies. Methods and techniques for the growth of bacterial cells, the introduction of isolated DNA molecules into host cells, and the isolation, cloning and sequencing of isolated nucleic acid molecules, etc., are a few examples of such methods and techniques. These methods and techniques are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods In Molecular Biology* (1986), J.H. Miller, *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1972); J.H. Miller, *A Short Course in Bacterial Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1992); M. Singer and P. Berg, *Genes & Genomes*, University Science Books, Mill Valley, California (1991); J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989); P.B. Kaufman *et al.*, *Handbook of Molecular and Cellular Methods in Biology and Medicine*, CRC Press, Boca Raton, Florida (1995); *Methods in Plant Molecular Biology and Biotechnology*, B.R. Glick and J.E. Thompson, eds., CRC Press, Boca Raton, Florida (1993); and P.F. Smith-Keary, *Molecular Genetics of Escherichia coli*, The Guilford Press, New York, NY (1989), all of which are incorporated herein by reference in their entireties.

Unless otherwise indicated, all nucleotide sequences newly described herein were determined using an automated DNA sequencer (such as the Model

373 from Applied Biosystems, Inc.). Therefore, as is known in the art, for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art.

In certain embodiments, polynucleotides of the invention comprise a nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a sequence selected from the group consisting of SEQ ID NO:17, SEQ ID NO:18; and SEQ ID NO:20, or a complementary sequence thereof.

By a polynucleotide comprising a nucleic acid, the sequence of which is at least, for example, 95% "identical" to a reference nucleotide sequence is intended that the nucleic acid sequence is identical to the reference sequence except that the nucleic acid sequence may include up to five mismatches per each 100 nucleotides of the reference nucleic acid sequence. In other words, to obtain a nucleic acid, the sequence of which is at least 95% identical to a reference nucleic acid sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The reference (query) sequence may be any one of the entire nucleotide sequences shown in SEQ ID NO:17, SEQ ID NO:18, or SEQ ID NO:20, or any fragment of any of these sequences, as described *infra*.

As a practical matter, whether any particular nucleic acid sequence is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to, for instance, a nucleotide sequence consisting of SEQ ID NO:17; SEQ ID NO:18, or SEQ ID NO:20, or a complementary sequence thereof, can be determined conventionally using sequence analysis computer programs such as a OMIGA® Version 2.0 for Windows, available from Oxford Molecular, Ltd. (Oxford, U.K.).

OMIGA uses the CLUSTAL W alignment algorithm using the slow full dynamic programming alignment method with default parameters of an open gap penalty of 10 and an extend gap penalty of 5.0, to find the best alignment between two nucleotide sequences. When using CLUSTAL W or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence such that gaps, mismatches, or insertions of up to 5% of the total number of nucleotides in the reference sequence are allowed. Other sequence analysis programs, known in the art, can be used in the practice of the invention.

This embodiment of the present invention is directed to polynucleotides comprising a nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence of SEQ ID NO:17, SEQ ID NO:18, and SEQ ID NO:20, or a complementary sequence thereof, irrespective of whether they have functional activity. This is because even where a particular polynucleotide does not have functional activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe, an S1 nuclease mapping probe, or a polymerase chain reaction (PCR) primer.

Preferred, however, are polynucleotides comprising a nucleic acid, the sequence of which is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence of SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:20, or a complementary sequence thereof, which do, in fact, have functional activity in *Corynebacterium* species.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of a polypeptide is intended that the amino acid sequence of the claimed polypeptide is identical to the reference sequence except that the claimed polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the polypeptide. In other words, to obtain a polypeptide having an amino



acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in SEQ ID NO:2 or to the amino acid sequence encoded by a nucleic acid sequence can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

In a specific embodiment, the identity between a reference sequence (query sequence, a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, is determined using the FASTDB computer program based on the algorithm of Brutlag *et al.* (Comp. App. Biosci. 6:237-245 (1990)). Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

According to this embodiment, if the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction is made to the results to take into consideration the fact that the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. A determination of whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of this embodiment. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence. For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not

manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of this embodiment.

### C. *Methods and Processes of the Invention*

Various embodiments of the invention provide methods to increase the production of an amino acid and processes for the production of an amino acid from a *Corynebacterium* species host cell. Particularly preferred *Corynebacterium* species of the methods and processes of the invention include: *Corynebacterium glutamicum*, *Brevibacterium flavum*, *Brevibacterium lactofermentum* and other *Corynebacteria* and *Brevibacteria* species known in the art.

As will be understood by those skilled in the art, the term “*Corynebacterium* species” includes those organisms previously identified in the literature as “*Brevibacterium* species,” for example *Brevibacterium flavum* and *Brevibacterium lactofermentum* which have now been reclassified into the genus *Corynebacterium* (*Int. J. Syst. Bacteriol.* 41: 255 (1981)).

Amino acid biosynthetic pathway genes embodied by the methods and processes described herein include those for L-glycine, L-alanine, L-methionine, L-phenylalanine, L-tryptophan, L-proline, L-serine, L-threonine, L-cysteine, L-tyrosine, L-asparagine, L-glutamine, L-aspartic acid, L-glutamic acid, L-lysine, L-arginine, L-histidine, L-isoleucine, L-leucine, and L-valine biosynthesis. Particularly preferred embodiments are drawn to biosynthetic pathway genes for L-lysine (Sahm *et al.*, *Ann. N. Y. Acad. Sci.* 782: 25-39 (1996)), L-threonine, L-isoleucine, L-tryptophan, and L-valine.

By way of example, the amino acid pathway for L-lysine biosynthesis is well known to skilled artisans of amino acid production in *Corynebacterium* species. Genes encoding the enzymes important for the conversion of L-aspartate to L-lysine include the *ask*, *asd*, *dapA*, *dapB*, *ddh* and *lysA* genes

(Figure 1). Thus, the invention provides herein for exemplary purposes only, specific embodiments utilizing L-lysine biosynthetic pathway genes. Other embodiments drawn to the use of biosynthetic pathway genes for the synthesis of other amino acids are also encompassed by the invention described herein.

5           The methods to increase the production of an amino acid and the processes for the production of an amino acid of the invention both utilize a step requiring the transformation of an isolated nucleic acid molecule into a *Corynebacterium* species host cell. As known to one skilled in the art, transformation of an isolated nucleic acid molecule into a host cell may be effected by electroporation, transduction or other methods. These methods are described in the many standard laboratory manuals referenced and incorporated herein.

10           The methods to increase the production of an amino acid and the processes for the production of an amino acid of the invention both utilize a step requiring amplification of at least one amino acid biosynthesis pathway gene. As known to one skilled in the art, the term amplification means increasing the number of a gene or genes of an amino acid biosynthetic pathway by any means known in the art. Particularly preferred means of amplification include: (1) the addition an isolated nucleic acid molecule comprising copies of a gene or genes of a biosynthetic pathway by insertion into the chromosome of a host cell, for example by homologous recombination, and (2) the addition an isolated nucleic acid molecule comprising copies of a gene or genes of a biosynthetic pathway into a host cell by way of a self-replicating, extra-chromosomal vector, for example, a plasmid.

20           Another method of the invention to increase the production of an amino acid comprises increasing the expression of at least one amino acid biosynthetic pathway gene. Preferred methods of increasing expression comprise using heterologous promoters, regulated promoters, unregulated promoters and combinations thereof.

25           Methods of inserting an isolated nucleic acid molecule into the chromosome of a host cell are known to those skilled in the art. For example,

insertion of isolated nucleic acid molecules into the chromosome of *Corynebacterium* species may be done utilizing the pK184 plasmid described by Jobling, M. *et al.*, *Nucleic Acids Research* 18(17): 5315-5316 (submitted 1990). Because these vectors lack a *Corynebacterium* species origin of replication and contain a selectable marker such as kanamycin (*kan*), cells will only be capable of growing under selection if the vector has been inserted into the host cell chromosome by homologous recombination.

In alternative embodiments, the invention also provides methods for increasing amino acid production and processes for the production of an amino acid wherein biosynthetic pathway gene amplification is accomplished through the introduction into a host cell of a self-replicating, extra-chromosomal vector, *e.g.*, a plasmid, comprising an isolated nucleic acid molecule encoding an amino acid biosynthetic pathway gene or genes. Suitable plasmids for these embodiments include pSR1 and other derivatives of pSR1 (Archer, J. *et al.*, *J. Gen. Microbiol.* 139: 1753-1759 (1993)).

For various embodiments of the invention drawn to a method to increase production of an amino acid, screening for increased production of an amino acid, for example L-lysine, may be determined by directly comparing the amount of L-lysine produced in culture by a *Corynebacterium* species host strain to that of a *Corynebacterium* species transformed host strain in which an amino acid biosynthesis gene or genes are amplified. The level of production of the amino acid of choice may conveniently be determined by the following formula to calculate the percent yield from dextrose:  $[(\text{g amino acid/L} / (\text{g dextrose consumed/L})) * 100]$ .

In one embodiment, the invention provides a method to increase the production of an amino acid comprising: (a) transforming a *Corynebacterium* species host cell with an isolated polynucleotide molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of SEQ ID NO:2; (b) amplifying the number of at least one of the biosynthetic pathway genes for said amino acid in the chromosome of said host cell;

(c) selecting a transformed host cell; and (d) screening for increased production of said amino acid from said transformed host cell relative to said host cell.

In a particularly preferred embodiment, the invention provides a method to increase the production of an amino acid comprising transforming a *Corynebacterium* species host cell with an isolated polynucleotide molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of SEQ ID NO:2; and further comprising at least one of the following: a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *asd* amino acid sequence; a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *dapA* amino acid sequence; a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *dapB* amino acid sequence; a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *ddh* amino acid sequence; a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *lysA* amino acid sequence; a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *lysA* amino acid sequence; and a nucleic acid molecule encoding a *Corynebacterium* species lysine pathway *ORF2* amino acid sequence.

In another particular embodiment of the method, the isolated polynucleotide molecule further comprises at least one of the following: a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another particular embodiment of the method, the isolated polynucleotide molecule further comprises the following: a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid

molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another particular embodiment of the method, the isolated polynucleotide molecule further comprises the following: a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another particular embodiment of the method, the isolated polynucleotide molecule further comprises the following: a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another particular embodiment of the method, the polynucleotide molecule further comprises the following: a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

In another embodiment of the method, the method further comprises growing said transformed host cell in a medium; and purifying an amino acid produced by said transformed host cell.

It is another object of the invention to provide an isolated polynucleotide molecule comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2; and at least one additional *Corynebacterium* species lysine pathway gene selected from the group consisting of a nucleic acid molecule encoding an *asd* polypeptide; a nucleic acid molecule encoding a *dapA* polypeptide; a nucleic acid molecule encoding a *dapB* polypeptide; a nucleic acid molecule encoding a *ddh* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; and a nucleic acid molecule encoding an *ORF2* polypeptide. In a preferred embodiment, said *asd* polypeptide is SEQ ID NO:4; said *dapA* polypeptide is SEQ ID NO:6; said *dapB* polypeptide is SEQ ID NO:8; said *ddh* polypeptide is SEQ ID NO:10; said *lysA* polypeptide is SEQ ID NO:21; said *lysA* polypeptide is SEQ ID NO:14; and said *ORF2* polypeptide is SEQ ID NO:16.

It is another object of the invention to provide an isolated polynucleotide molecule comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO 2; a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

It is another object of the invention to provide an isolated polynucleotide molecule comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO: 2; a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid



sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

It is another object of the invention to provide an isolated polynucleotide molecule comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2; a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

It is another object of the invention to provide an isolated polynucleotide molecule comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2; a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4; a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6; a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

It is a further object of the invention to provide an isolated polynucleotide molecule comprising pK184-KDAB. It is a further object of the invention to provide an isolated polynucleotide molecule comprising pK184-KDABH'L. It is a further object of the invention to provide an isolated polynucleotide molecule comprising pD11-KDABH'L. It is a further object of the invention to provide an isolated polynucleotide molecule comprising pD2-KDABHL.

It is a further object of the invention to provide a vector comprising the isolated polynucleotide molecule comprising a nucleotide sequence encoding a

polypeptide comprising the amino acid sequence of SEQ ID NO 2; and further comprising at least one additional *Corynebacterium* species lysine pathway gene selected from the group consisting of a nucleic acid molecule encoding an *asd* polypeptide; a nucleic acid molecule encoding a *dapA* polypeptide; a nucleic acid molecule encoding a *dapB* polypeptide; a nucleic acid molecule encoding a *ddh* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; and a nucleic acid molecule encoding an *ORF2* polypeptide.

It is a further object to provide a host cell comprising a vector comprising the isolated polynucleotide molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of SEQ ID NO 2; and further comprising at least one additional *Corynebacterium* species lysine pathway gene selected from the group consisting of a nucleic acid molecule encoding an *asd* polypeptide; a nucleic acid molecule encoding a *dapA* polypeptide; a nucleic acid molecule encoding a *dapB* polypeptide; a nucleic acid molecule encoding a *ddh* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; a nucleic acid molecule encoding a *lysA* polypeptide; and a nucleic acid molecule encoding an *ORF2* polypeptide.

It is a further object to provide a host cell wherein said host cell is a *Brevibacterium* selected from the group consisting of *Brevibacterium flavum* NRRL-B30218, *Brevibacterium flavum* NRRL-B30219, *Brevibacterium lactofermentum* NRRL-B30220, *Brevibacterium lactofermentum* NRRL-B30221, *Brevibacterium lactofermentum* NRRL-B30222, *Brevibacterium flavum* NRRL-30234 and *Brevibacterium lactofermentum* NRRL-30235. In another embodiment, the host cell is *Escherichia coli* DH5  $\alpha$  MCR NRRL-B30228. In another embodiment, the host cell is a *C. glutamicum* selected from the group consisting of *C. glutamicum* NRRL-B30236 and *C. glutamicum* NRRL-B30237.

The invention provides processes for the production of an amino acid. In one embodiment, the invention provides a process for producing an amino acid comprising: (a) transforming a *Corynebacterium* species host cell with an isolated nucleic acid molecule; (b) amplifying the number of chromosomal

copies of at least one of the biosynthetic pathway genes for said amino acid; (c) selecting a transformed host cell; (d) growing said transformed cell in a medium; and (e) purifying said amino acid.

The invention is also directed to an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 19. In one embodiment of the invention, the polypeptide has at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 19. The invention is also directed to an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO: 19. In one embodiment, the isolated polynucleotide comprises a nucleic acid having the sequence of SEQ ID NO: 18.

The invention is also directed to a vector comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO: 19. In one embodiment, the invention is directed to a host cell comprising a vector encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 19. In one embodiment, the host cell is NRRL B30360.

The invention is also directed to a method comprising transforming a *Corynebacterium* species host cell with the polynucleotide molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 19, and selecting a transformed host cell. In one embodiment, the method further comprises screening for increased amino acid production. In a preferred embodiment, the amino acid screened for is lysine. In one embodiment, the polynucleotide molecule is integrated into said host cell's chromosome, thereby increasing the total number of said amino acid biosynthetic pathway genes in said host cell chromosome.

In another embodiment, the polynucleotide molecule further comprises at least one of the following: (a) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ask* amino acid sequence; (b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *asd* amino acid sequence; (c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapA* amino acid sequence; (d) a nucleic acid molecule encoding

a *Corynebacterium species* lysine pathway *dapB* amino acid sequence; (e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddh* amino acid sequence; (f) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence; (g) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence; and, (h) a nucleic acid molecule encoding an ORF2 polypeptide having SEQ ID NO:16. In this embodiment, the method further comprises screening for increased amino acid production. In another embodiment, the amino acid screened for is lysine.

In another embodiment of the method, the polynucleotide molecule further comprises: (a) a nucleic acid molecule encoding the *ask* amino acid sequence having SEQ ID NO:2; (b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *asd* amino acid sequence; (c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapB* amino acid sequence; (d) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddh* amino acid sequence; and, (e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence. In one embodiment of this method, the method further comprises screening for increased amino acid production.

The invention is also directed to an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:21. In one embodiment, the polypeptide has at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21. The invention also comprises an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21. The invention is further comprises a polynucleotide molecule comprising a nucleic acid having the sequence of SEQ ID NO:20. In one embodiment the invention comprises a vector comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21. The invention further comprises a host cell

comprising the vector comprising the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21.

5 In one embodiment, the invention comprises a host cell selected from the group consisting of NRRL B30218, NRRL B30220 and NRRL B30222.

The invention is further directed to a method comprising transforming a *Corynebacterium* species host cell with a polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 21, and selecting a transformed host cell. The method further comprises screening for increased amino acid production; in particular, for lysine production. In one embodiment, the polynucleotide molecule is integrated into said host cell's chromosome, thereby increasing the total number of said amino acid biosynthetic pathway genes in said host cell chromosome. In one embodiment the method further comprises a polynucleotide molecule further comprising at least one of the following: (a) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ask* amino acid sequence; (b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ask* amino acid sequence having SEQ ID NO. 2; (c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *asd* amino acid sequence; (d) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapA* amino acid sequence; (e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapB* amino acid sequence; (f) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddlh* amino acid sequence; (g) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ORF2* amino acid sequence; and, (h) a nucleic acid molecule encoding a truncated *Corynebacterium species* lysine pathway *ORF2* amino acid sequence. In one embodiment, the method further comprises screening for increased amino acid production. In another embodiment, the amino acid screened for is lysine.

Another embodiment of the invention is also directed to an isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polynucleotide molecule further comprises a promoter sequence having SEQ ID NO:17. In one embodiment, the promoter sequence has at least 95% sequence identity to SEQ ID NO:17. In one embodiment, the promoter sequence having at least 95% sequence identity to SEQ ID NO:17 is operably directly linked to the LysA gene. In another embodiment of the invention, there is a vector comprising the isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polynucleotide molecule further comprises a promoter sequence wherein said promoter sequence has at least 95% sequence identity to SEQ ID NO:17. In another aspect of the invention, there is a host cell comprising the vector comprising the isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polynucleotide molecule further comprises a promoter sequence having at least 95% sequence identity to SEQ ID NO:17. In one embodiment, the host cell is NRRL B30359.

The invention is also directed to a method comprising transforming a *Corynebacterium* species host cell with the polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the polynucleotide molecule further comprises a promoter sequence having at least 95% sequence identity to SEQ ID NO:17, and selecting a transformed host cell. In one embodiment, the method further comprises screening for increased amino acid production. In another embodiment, the amino acid screened for is lysine. In another embodiment of the method, the polynucleotide molecule is integrated into said host cell's chromosome, thereby increasing the total number of amino acid biosynthetic pathway genes in said host cell chromosome. In another embodiment of the method, the polynucleotide molecule further comprises at least one of the following: (a) a nucleic acid molecule encoding a *Corynebacterium species*

lysine pathway *asd* amino acid sequence; (b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapA* amino acid sequence; (c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapB* amino acid sequence; (d) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddh* amino acid sequence; (e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ORF2* amino acid sequence; (f) a nucleic acid molecule encoding a truncated *Corynebacterium species* lysine pathway *ORF2* amino acid sequence; (g) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence; and, (h) a nucleic acid molecule encoding a truncated *Corynebacterium species* lysine pathway *lysA* amino acid sequence. In this embodiment, the method further comprises screening for increased amino acid production; in particular, for lysine production.

In a different embodiment of the method, the polynucleotide molecule comprises: (a) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *asd* amino acid sequence; (b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapA* amino acid sequence; (c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapB* amino acid sequence; (d) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddh* amino acid sequence; (e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ORF2* amino acid sequence; and, (f) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence. In this embodiment, the method further comprises screening for increased amino acid production. In a preferred embodiment, the amino acid is lysine.

A variety of media known to those skilled in the art may be used to support cell growth for the production of an amino acid. Illustrative examples of suitable carbon sources include, but are not limited to: carbohydrates, such as glucose, fructose, sucrose, starch hydrolysate, cellulose hydrolysate and molasses; organic acids, such as acetic acid, propionic acid, formic acid, malic acid, citric acid, and fumaric acid; and alcohols, such as glycerol. Illustrative

examples of suitable nitrogen sources include, but are not limited to: ammonia, including ammonia gas and aqueous ammonia; ammonium salts of inorganic or organic acids, such as ammonium chloride, ammonium phosphate, ammonium sulfate and ammonium acetate; and other nitrogen-containing sources, including meat extract, peptone, corn steep liquor, casein hydrolysate, soybean cake hydrolysate, urea and yeast extract.

A variety of fermentation techniques are known in the art which may be employed in processes of the invention drawn to the production of amino acids. Generally, amino acids may be commercially produced from the invention in fermentation processes such as the batch type or of the fed-batch type. In batch type fermentations, all nutrients are added at the beginning of the fermentation. In fed-batch or extended fed-batch type fermentations one or a number of nutrients are continuously supplied to the culture, right from the beginning of the fermentation or after the culture has reached a certain age, or when the nutrient(s) which are fed were exhausted from the culture fluid. A variant of the extended batch of fed-batch type fermentation is the repeated fed-batch or fill-and-draw fermentation, where part of the contents of the fermenter is removed at some time, for instance when the fermenter is full, while feeding of a nutrient is continued. In this way a fermentation can be extended for a longer time.

Another type of fermentation, the continuous fermentation or chemostat culture, uses continuous feeding of a complete medium, while culture fluid is continuously or semi-continuously withdrawn in such a way that the volume of the broth in the fermenter remains approximately constant. A continuous fermentation can in principle be maintained for an infinite time.

In a batch fermentation an organism grows until one of the essential nutrients in the medium becomes exhausted, or until fermentation conditions become unfavorable (*e.g.*, the pH decreases to a value inhibitory for microbial growth). In fed-batch fermentations measures are normally taken to maintain favorable growth conditions, *e.g.*, by using pH control, and exhaustion of one or more essential nutrients is prevented by feeding these nutrient(s) to the culture. The microorganism will continue to grow, at a growth rate dictated by the rate



of nutrient feed. Generally a single nutrient, very often the carbon source, will become limiting for growth. The same principle applies for a continuous fermentation, usually one nutrient in the medium feed is limiting, all other nutrients are in excess. The limiting nutrient will be present in the culture fluid at a very low concentration, often unmeasurably low. Different types of nutrient limitation can be employed. Carbon source limitation is most often used. Other examples are limitation by the nitrogen source, limitation by oxygen, limitation by a specific nutrient such as a vitamin or an amino acid (in case the microorganism is auxotrophic for such a compound), limitation by sulphur and limitation by phosphorous.

The amino acid may be recovered by any method known in the art. Exemplary procedures are provided in the following: Van Walsem, H.J. & Thompson, M.C., *J. Biotechnol.* 59:127-132 (1997), and U.S. Pat. No. 3,565,951, both of which are incorporated herein by reference.

The invention described herein provides isolated nucleic acid molecules comprising at least one L-lysine amino acid biosynthesis gene. Unless otherwise indicated, all nucleotide sequences described herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules described herein were predicted by translation of the relative DNA sequence. Therefore, as is known in the art, for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art.

As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely

different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

The invention provides several isolated nucleic acid molecules encoding comprising at least one L-lysine amino acid biosynthesis pathway gene of *Corynebacterium glutamicum*. More specifically, the invention provides the following isolated nucleic acid molecules: the nucleotide sequence of the *ask* gene from the strain ATCC 21529 (SEQ ID NO:1); the nucleotide sequence of the *asd* gene from the strain ATCC 21529 (SEQ ID NO:3); the nucleotide sequence of the *dapA* gene from the strain NRRL-B11474 (SEQ ID NO:5); the nucleotide sequence of the *dapB* gene from the strain NRRL-B11474 (SEQ ID NO:7); the nucleotide sequence of the *ddh* gene from the strain NRRL-B11474 (SEQ ID NO:9) and the nucleotide sequence of the *ORF2* gene from the strain NRRL-B11474 (SEQ ID NO:15). In addition, also provided herein is the nucleotide sequence of *lysA* (SEQ ID NO:13) gene from plasmid pRS6 (Marcel, T., et al., *Molecular Microbiology* 4: 1819-1830 (1990)).

It is known in the art that amino acids are encoded at the nucleic acid level by one or more codons (code degeneracy). It is also known in the art that choice of codons may influence expression of a particular amino acid sequence (protein, polypeptide, etc.). Thus, the invention is further directed to nucleic acid molecules encoding the *ask* amino acid sequence of SEQ ID NO:2 wherein the nucleic acid molecule comprises any codon known to encode a particular amino acid. The invention is also further directed to nucleic acid sequences (SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 18 and 20) which comprise alternative codons in order to optimize expression of the protein or polypeptide.

In addition to the above described isolated nucleic acid molecules, the invention also provides isolated nucleic acid molecules comprising more than one L-lysine *Corynebacterium glutamicum* biosynthesis gene. Such isolated nucleic acid molecules are referred to as "cassette" constructs. These cassette constructs simplify for the practitioner the number of recombinant DNA manipulations required to achieve gene amplification of L-lysine biosynthesis genes.

In one embodiment drawn to a cassette construct, the invention provides an isolated nucleic acid molecule comprising: (a) a polynucleotide encoding the *Corynebacterium glutamicum* L-lysine pathway *ask* amino acid sequence of SEQ ID NO:2; and (b) at least one additional *Corynebacterium* species L-lysine pathway gene selected from the group consisting of: (1) a polynucleotide encoding the *asd* polypeptide; (2) a polynucleotide encoding the *dapA* polypeptide; (3) a polynucleotide encoding the *dapB* polypeptide; (4) a polynucleotide encoding the *ddh* polypeptide; (5) a polynucleotide encoding the *lysA* polypeptide, and (6) a polynucleotide encoding the *ORF2* polypeptide.

The isolated nucleic acid molecules of the invention are preferably propagated and maintained in an appropriate nucleic acid vector. Methods for the isolation and cloning of the isolated nucleic acid molecules of the invention are well known to those skilled in the art of recombinant DNA technology. Appropriate vectors and methods for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, N.Y., 1989, the disclosure of which is hereby incorporated by reference.

A great variety of vectors can be used in the invention. Such vectors include chromosomal, episomal and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids and from bacteriophage, as well as vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used in accordance with this aspect of the present invention. Generally, any vector suitable to maintain and propagate a polynucleotide in a bacterial host may be used in this regard.

A large numbers of suitable vectors and promoters for use in bacteria are known, many of which are commercially available. Preferred prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (such as, for example, pBR322, ColEI, pSC101, pACYC 184,  $\pi$ VX). Such plasmids are, for example, disclosed by Maniatis, T., *et al.*, *In: Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1982)).

The following vectors are provided by way of example: pET (Novagen), pQE70, pQE60, pQE-9 (Qiagen), pBs, phagescript, psiXI74, pBlueScript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene), pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia).

5 Preferred vectors for the isolated nucleic acid molecules of the invention include the pFC1 to pFC7 novel family of combinatorial cloning vectors (Lonsdale, D.M., *et al.*, *Plant Molecular Biology Reporter* 13: 343-345 (1995)), the pK184 vector (Jobling, M.G. and Homes, R.K., *Nucleic Acid Research* 18: 5315-5316 (1990)).

10 Another group of preferred vectors are those that are capable of autonomous replication in *Corynebacterium* species. Such vectors are well known to those skilled in the art of amino acid production by way of microbial fermentation, examples of which include pSR1, pMF1014 $\alpha$  and vectors derived therefrom.

15 The invention provides an isolated amino acid sequence of the *ask* polypeptide of the strain ATCC 21529 (SEQ ID NO:2). The isolated *ask* amino sequence disclosed herein possesses unique properties with respect to feedback resistance of *ask* enzyme activity to accumulated levels of L-lysine and L-threonine in the culture medium. When compared to the DNA sequences of  
20 other *Corynebacterium glutamicum ask-asd* gene sequences, the invention discloses a threonine to isoleucine change at amino acid residue 380 which results in resistance to feedback inhibition. The invention also includes other amino acid changes at residue 380 which result in decreased *ask* enzyme sensitivity to L-threonine and/or L-lysine.

25 In addition, and as described in more detail herein, the vector may contain control regions that regulate as well as engender expression. Generally, such regions will operate by controlling transcription, such as inducer or repressor binding sites and enhancers, among others.

30 Vectors of the present invention generally will include a selectable marker. Such markers also may be suitable for amplification or the vectors may contain additional markers for this purpose. In this regard, vectors preferably

contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Such markers include, but are not limited to, an antibiotic resistance gene such as a chloramphenicol, ampicillin, or kanamycin resistance gene, or an autotrophic gene which allows the host cell to grow in the absence of a nutrient for which the host cell strain is normally auxotrophic.

If the vector is intended to be maintained in the host cell extrachromosomally, it will contain, in addition an origin of replication which will allow it to replicate in the *Corynebacterium* species host cell. Alternatively, if it is desired that the vector integrate into the *Corynebacterium* species chromosome, the vector is constructed such that it cannot replicate in *Corynebacterium*. For example, such a vector might be capable of propagation in another organism, for example, *E. coli*, but lack the proper origin of replication to be propagated in *Corynebacterium*. In another aspect of this embodiment, the vector is a shuttle vector which can replicate and be maintained in more than one host cell species, for example, such a shuttle vector might be capable of replication in a *Corynebacterium* host cell such as a *C. glutamicum* host cell, and also in an *E. coli* host cell.

The invention further provides the following isolated the amino acid sequences: the amino acid sequence of the *asd* polypeptide of the strain ATCC 21529 (SEQ ID NO:4); the amino acid sequence of the *dapA* polypeptide of the strain NRRL-B11474 (SEQ ID NO:6); the amino acid sequence of the *dapB* polypeptide of the strain NRRL-B11474 (SEQ ID NO:8); the amino acid sequence of the *ddh* polypeptide of the strain NRRL-B11474 (SEQ ID NO:10) and the amino acid sequence of the *ORF2* polypeptide of the strain NRRL-B11474 (SEQ ID NO:16). In addition, also provided herein is the amino acid sequence of *lysA* (pRS6) (Marcel, T., *et al.*, *Mol. Microbiol.* 4: 819-830 (1990)) (SEQ ID NO:14).

In addition to the isolated polypeptide sequences defined by the specific sequence disclosures disclosed above, the invention also provides the amino acid sequences encoded by the deposited clones.

It will be recognized in the art that some amino acid sequences of the invention can be varied without significant effect of the structure or function of the proteins disclosed herein. Variants included may constitute deletions, insertions, inversions, repeats, and type substitutions so long as enzyme activity is not significantly affected. Guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J.U., *et al.*, "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science* 247:1306-1310 (1990).

The strains of the invention may be prepared by any of the methods and techniques known and available to those skilled in the art. Introduction of gene constructs of the invention into the host cell can be effected by electroporation, transduction or other methods. These methods are described in the many standard laboratory manuals referenced and incorporated herein.

Various embodiments of the invention provide strains with increased L-lysine production as a result of gene amplification. By gene amplification is meant increasing the number of copies above the normal single copy number of an L-lysine biosynthesis pathway gene by a factor of 2, 3, 4, 5, 10, or more copies.

In one embodiment of the invention, the additional copies of the L-lysine biosynthesis pathway gene(s) may be integrated into the chromosome. Another embodiment of the invention provides that the additional copies of the L-lysine biosynthesis pathway gene(s) are carried extra-chromosomally. Amplifications by a factor of 5 or less may be obtained by introducing the additional gene copies into the chromosome of the host strain by way of single event homologous recombination. In a most preferred embodiment, the recombination event results in the introduction of one additional copy of the copy of the gene or genes of interest. If more than 5 copies of the genes are desired, then the invention also provides for the use of multicopy plasmids carrying the recombinant DNA construct of the invention.

Representative examples of appropriate hosts for isolated nucleic acid molecules of the invention include, but are not limited to, bacterial cells, such as

*C. glutamicum*, *Escherichia coli*, *Streptomyces* and *Salmonella typhimurium* cells; and fungal cells, such as yeast cells. Appropriate culture media and conditions for the above-described host cells are known in the art.

Particularly preferred host cells of the invention include:  
5 *Corynebacterium glutamicum*, *Brevibacterium flavum* and *Brevibacterium lactofermentum*.

Applicants have deposited clones carrying the pK184-KDABH'L multi-gene constructs at an acceptable International Depositary Authority in accordance with the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. The deposits have been made with the Agricultural Research Service, Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604. Deposits made in which the pK184-KDAB or pK184-KDABH'L multi-gene constructs have been integrated into the chromosome of a host cell include the following: (1) the pK184-KDAB plasmid, integrated into the chromosome, deposited as NRRL-B30219 and NRRL-B30221 and (2) the pK184-KDABH'L plasmid, integrated into the chromosome, deposited as NRRL-B30218, NRRL-B30220, and NRRL-B30222. In addition, the pK184-KDABH'L multigene construct in a plasmid configuration, carried in *E. coli* DH5 $\alpha$  MCR, was deposited as NRRL-B30228.  
10 The six gene construct (pDElia2-KDABHL) was deposited in *E. coli* (NRRL-B30233). *C. glutamicum* comprising pK184-KDABH'L was deposited as NRRL-B30236. *C. glutamicum* comprising pK184-KDABHL was deposited as NRRL-B30237. *Brevibacterium flavum* comprising pDElia2-KDABHL was deposited as NRRL-B30234. *Brevibacterium lactofermentum* comprising  
15 pDElia2-KDABHL was deposited as NRRL-B30235.

It is an object of the invention to provide a method of producing lysine comprising culturing the host cells comprising the amino acid sequence of SEQ ID NO:2 wherein said host cells comprise one or more of: (a) increased enzyme activity of one or more lysine biosynthetic pathway enzymes compared to the genetically unaltered host cell; (b) one or more copies of each gene encoding a  
20 lysine biosynthetic pathway enzyme; and, (c) alteration of one or more

transcription factors regulating transcription of one or more genes encoding a lysine biosynthetic pathway enzyme, wherein said host cell produces lysine in said culture medium. In one embodiment of the method, said increased enzyme activity comprises overexpressing one or more genes encoding one or more lysine biosynthetic pathway enzymes. In one embodiment of the method, said one or more genes are operably linked directly or indirectly to one or more promoter sequences. In another embodiment of the method, said operably linked promoter

sequences are heterologous, endogenous, or hybrid. In a preferred embodiment of the method, said promoter sequences are one or more of: a promoter sequence from the 5' end of genes endogenous to *C. glutamicum*, a promoter sequence from plasmids that replicate in *C. glutamicum*, and, a promoter sequence from the genome of phage which infect *C. glutamicum*. In a preferred embodiment of the method, one or more of said promoter sequences are modified. In another preferred embodiment, said modification comprises truncation at the 5' end, truncation at the 3' end, non-terminal insertion of one or more nucleotides, non-terminal deletion of one or more nucleotides, addition of one or more nucleotides at the 5' end, addition of one or more nucleotides at the 3' end, and, combinations thereof.

In another embodiment of the method, said increased enzyme activity results from the activity of one or more modified lysine biosynthetic pathway enzymes wherein said enzyme modification results in a change in kinetic parameters, allosteric regulation, or both, compared to the enzyme lacking the modification. In one embodiment of the method, said change in kinetic parameters is a change in  $K_m$ ,  $V_{max}$  or both. In another embodiment of the method, said change in allosteric regulation is a change in one or more enzyme allosteric regulatory sites. In one embodiment, said change in allosteric regulation is a change in the affinity of one or more enzyme allosteric regulatory sites for the ligand or ligands. The ligands may be the same or different. In one embodiment, said enzyme modification is a result of a change in the nucleotide sequence encoding said enzyme. In one embodiment, said change in said



nucleotide sequence is an addition, insertion, deletion, substitution, or a combination thereof, of one or more nucleotides.

In another embodiment of the method, said alteration of one or more transcription factors comprises one or more mutations in transcription inhibitor proteins, one or more mutations in transcription activator proteins, or both, wherein said one or more mutations increases transcription of the target nucleotide sequence compared to the transcription by said one or more transcription factors lacking said alteration. In one embodiment, said one or more mutations is a change in said nucleotide sequence encoding said transcription factor. In another embodiment, said change in said nucleotide sequence is an addition, insertion, deletion, substitution, or a combination thereof, of one or more nucleotides.

All patents and publications referred to herein are expressly incorporated by reference in their entirety.

## ***Examples***

### ***Example 1***

#### ***Preparation of L-Lysine Pathway Multi-gene Constructs pK184-KDAB and pK184-KDABH'L***

Applicants have created L-lysine amino acid biosynthetic pathway multi-gene constructs for the purpose of amplifying the number of one or more of the genes of this pathway in the chromosome of *Corynebacterium* species. Also, through careful study of the L-lysine biosynthesis genes of strain ATCC 21529, Applicants have identified an amino acid change of threonine to isoleucine at amino acid residue 380 of the *ask* gene of ATCC 21529. Compared to the DNA sequences of other *Corynebacterium glutamicum ask* genes, a threonine to isoleucine change at amino acid residue 380 was observed (Figure 19), which is responsible for the unusual feedback resistant properties with respect to aspartate kinase enzyme regulation.

The isolated nucleic acid molecules encoding L-lysine, amino acid biosynthesis pathway genes utilized in the present invention are from the following sources:

Gene(s)	Source
<i>ask-asd</i>	Strain ATCC 21529;
<i>dapA</i>	Strain NRRL B11474;
<i>dapB</i>	Strain NRRL B11474;
<i>ddh</i>	Strain NRRL B11474;
<i>lysA</i>	Plasmid pRS6 (Marcel, T., <i>et al.</i> , <i>Mol. Microbiol.</i> 4: 819-830 (1990)) carrying the <i>lysA</i> gene isolated from strain AS019, which was derived from ATCC 13059;
<i>lysA</i>	NRRL B11474;
<i>lysA</i>	NRRL B11474 (full length); and,
<i>ORF2</i>	Strain NRRL B11474.

As one skilled in the art would know, the invention is not limited to the specific strain origins that Applicants present for the isolated nucleic acid molecules of the invention. Any strain of *Corynebacterium* species, particularly that of *Corynebacterium glutamicum*, may be utilized for the isolation of nucleic acid molecules that will be used to amplify the number of chromosomally located amino acid biosynthetic pathway genes. Particularly preferred strains include: NRRL-B11474, ATCC 21799, ATCC 21529, ATCC 21543, and E12.

Methods and techniques common to the art of recombinant DNA technology were used in making the multi-gene constructs of the invention, as may be found in the many laboratory manuals cited and incorporated herein, for example as found in J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

The polymerase chain reaction (PCR) technique is used extensively in the making of the multi-gene constructs of the invention. In a typical reaction, the

standard 10X stock solution (100 mM Tris-HCL, pH 8.3, 500 mM KCL, 1.5 mM MgCl<sub>2</sub>) is diluted to 1X for use. Typical reaction conditions were used for PCR amplification: 10 mM Tris, pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.01% gelatin, 200  $\mu$ M deoxynucleotides, 0.2-1.0  $\mu$ M primers and 2.5 U/100  $\mu$ l pfu polymerase. Standard cycling parameters were also employed in PCR reactions: For 30 cycles, template denaturation was performed at 94 °C for 1 min; 55 °C annealing temperature was performed for 1 min (or annealing temperature appropriate for particular primer pair); product extension was performed at 72 °C for 1 min (if product is <500 bp), 3 min (if product is >500 bp); and at the end of cycling, a final extension at 72 °C for 7 min was performed.

The primers utilized for cloning experiments included:

*ask*: 5'-GGGTACCTCGCGAAGTAGCACCTGTAC-3';

*asd*: 5'-GCGGATCCCCATCGCCCCTCAAAGA-3';

*dapB*: 5'-AACGGGCGGTGAAGGGCAACT-3';

*dapA*: 5'-TGAAAGACAGGGGTATCCAGA-3';

*ddh* 5'-CCATGGTACCAAGTGC GTGGCGAG-3';

5'-CCATGGTACCACACTGTTTCCTTGC-3';

*argS*: 5'-CTGGTTCCGGCGAGTGGAGCCGACCATTCCGCGAGG-3'; and

*lysA*: 5'-CTCGCTCCGGCGAGGTCTGGAGGCAACTTCTGCGACG-3', a

primer that anneals internally to *lysA* (about 500bp upstream to the end of *lysA*).

'*LysA*' is a truncated form obtained from *lysA*.

Applicants utilized standard PCR and subcloning procedures in cloning the coding regions of *ask-asd*, *dapB-ORF2-dapA*, *ddh*, '*lysA*', and *lysA*. Construction procedures and intermediate plasmids are described in Figure 18. Applicants performed the following steps (Figure 18) in constructing the following vectors used in the L-lysine biosynthetic pathway:

1. pGEM-*ask-asd*: an approximately 2.6 Kb PCR product containing the *ask-asd* operon of ATCC21529 using primers *ask* and *asd* was cloned into pGEM-T (Promega pGEM-T vector systems);

2. pADM21: an approximately 1.3Kb PCR product (with an engineered KpnI site on both primers) of NRRL-B11474 *ddh* coding region was cloned into pADM20;

3. pUC 1 8-*ddh*: an approximately 1.3Kb KpnI fragment of pADM21 containing *ddh* (NRRL-B11474) was subcloned into pUC 18 at the KpnI site;

4. pLIC 1.7-*argS*-*lysA*: PCR product using template NRRL-B11474 genomic DNA and primers *argS* and *lysA* was cloned into pPMG-LIC cloning vector (PharMingen);

5. pM4-*dapB*-*ORF2-dapA*:: an approximately 3 Kb PCR product using primers *dapB* and *dapA* was cloned into pM4 at the XbaI site;

6. pFC3-*ask-asd*: an approximately 2.6 Kb NsiI-ApaI fragment of pGEMT-*ask-asd* was cloned into pFC3 cut with PstI and ApaI;

7. pFC1-*ddh*: ~1.3 Kb SalI-EcoRI fragment of pUC18-*ddh* was cloned into pFC1 cut with SalI and EcoRI;

8. pFC1-*ddh*-*lysA*: an approximately 1.5 Kb EcoRI fragment (containing the truncated *lysA* DNA) of pLIC1.7-*argS*-*lysA* was cloned into pFC1-*ddh* at the EcoRI site;

9. pFC5-*dapB*-*ORF2-dapA*: an approximately 3.4 Kb BamHI-BgIII fragment of pM4-*dapB*-*ORF2-dapA* was cloned into pFC5 at the BamHI site;

10. pFC5-*dapB*-*ORF2-dapA-ddh*-*lysA*: ~2.8 Kb NheI fragment of pFC1-*ddh*-*lysA* was cloned into pFC5-*dapB*-*ORF2-dapA* at the NheI site;

11. pFC3-*ask-asd-dapB*-*ORF2-dapA-ddh*-*lysA*: ~6.2 Kb NotI fragment of pFC5-*dapB*-*ORF2-dapA-ddh*-*lysA* was cloned into pFC3-*ask-asd* at the NotI site;

12. pDElia9-*ask-asd-dapB*-*ORF2-dapA-ddh*-*lysA* (pDElia9-KDABH'L): ~8.8 Kb PmeI fragment of pFC3-*ask-asd-dapB*-*ORF2-dapA-ddh*-*lysA* was cloned into pDElia9 at the EcoRV site; and

13. pK184-*ask-asd-dapB*-*ORF2-dapA-ddh*-*lysA* (pK184-KDABH'L): an approximately 8.8 Kb PmeI fragment of pFC3-*ask-asd-dapB*-*ORF2-dapA-ddh*-*lysA* was cloned into pK184 at the HincII or SmaI site.

14. pFC5-*ask-asd-dapB-ORF2-dapA* (pFC5-KDAB): ~2.6 Kb KpnI-SmaI fragment of pFC3-*ask-asd* was cloned into pFC5-*dapB-ORF2-dapA* cut with KpnI and SmaI.

5 15. pK184-*ask-asd-dapB-ORF2-dapA* (pK184-KDAB): ~7 Kb KpnI-PmeI fragment of pFC5-*ask-asd-dapB-ORF2-dapA* was cloned into pK184 cut with KpnI and HincII.

Thus, Applicants have made the following L-lysine multi-gene constructs:

10 1. pK184-KDABH'L, wherein "K" represents a nucleotide sequence encoding the *ask* polypeptide; "D" represents a nucleotide sequence encoding the *asd* polypeptide; "A" represents a nucleotide sequence encoding the *dapA* polypeptide; "B" represents a nucleotide sequence encoding the *dapB* polypeptide; "H'" represents a nucleotide sequence encoding the *ddh* polypeptide; and "'L" represents a nucleotide sequence encoding part of the '*lysA*' polypeptide. This construct is referred to as a truncated 6 gene construct. The  
15 pK184-KDABHL construct, constructed *infra*, is referred to as a full length 6 gene construct.

2. pK184-KDAB, wherein "K" represents a nucleotide sequence encoding the *ask* polypeptide; "D" represents a nucleotide sequence encoding the *asd* polypeptide; "A" represents a nucleotide sequence encoding the *dapA* polypeptide; and "B" represents a nucleotide sequence encoding the *dapB* polypeptide. This construct is referred to as a 4 gene construct.

Both pK184-KDABH'L and pK184-KDAB, as do the other constructs discussed herein, comprise the nucleotide sequence encoding the *ORF2* polypeptide.

25 It should be noted that in addition to the indicated polypeptide sequences encoded by the isolated nucleic acid sequences represented by "K", "D", "A", "B," "H," "L" and "'L", these isolated nucleic acid sequences also include native promoter elements for the operons represented therein. Thus, the *ask-asd* sequences have been cloned in a fashion that includes the respective native  
30 promoter elements; the *dapA* and *dapB* sequences, representing the operon *dapB-ORF2-dapA*, have been cloned in a fashion that includes the respective promoter

elements; the *ddh* sequence has been cloned in a fashion that includes the respective native promoter elements, and the *lysA* and *lysA* sequences have been cloned in a fashion that includes a native promoter element.

Alternative gene promoter elements may be utilized in the constructs of the invention. For example, known bacterial promoters suitable for this use in the present invention include the *E. coli lacI* and *lacZ* promoters, the *T3* and *T7* promoters, the *gpt* promoter, the lambda *PR* and *PL* promoters, the *trp* promoter, or promoters endogenous to the bacterial cells of the present invention. Other promoters useful in the invention include regulated promoters, unregulated promoters and heterologous promoters. Many such promoters are known to one of skill in the art. See Sambrook, E.F. *et al.*, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

### ***Example 2***

#### ***Two-Fold Amplification of L-lysine Amino Acid Biosynthesis Pathway Genes***

For exemplary purposes only, Applicants provide herein an example wherein at least one L-lysine amino acid biosynthesis pathway gene is amplified by a factor of 2 by way of (a) the introduction of an isolated nucleic acid molecule into a *Corynebacterium glutamicum* host cell, and (b) the subsequent single crossover homologous recombination event introducing said isolated nucleic acid molecule into said *Corynebacterium glutamicum* host cell chromosome.

As will be understood by those in the art, at least one or two or three or four or five or six or seven or eight or nine or ten or more amino acid biosynthesis pathway genes may be amplified, *i.e.*, increased in number, by a factor of at least one or two or three or four or five or six or seven or eight or nine or ten fold with minor variations of the example presented herein.

pK184-KDAB, pK184-KDABH'L and pD2-KDABHL(a full length 6 gene construct constructed in Example 4) plasmids were used in the construction of high yield derivative cell lines of the invention. This was accomplished by way of introducing plasmid pK184-KDAB, pK184-KDABH'L and pD2-KDABHL DNAs into a *Corynebacterium* species resulting in incorporation of pK184-KDAB, pK184-KDABH'L or pD2-KDABHL into the host cell chromosome via a single crossover homologous recombination event. Amplification of the amino acid biosynthetic pathway genes by way of chromosomal integration of the plasmid constructs of the invention provided increased L-lysine production in several *Corynebacterium* species strains.

For cell transformation experiments with the isolated nucleic acid molecules of the invention, the growth and preparation of competent cells may be done according to the following procedure: (1) picking a fresh, single colony of *Corynebacterium glutamicum* and growing a culture overnight in 10 mL CM (SM1) in a 250 mL shake flask at 30 degrees Celsius with agitation; (2) inoculating 200 mL of "Growth Media" with the overnight culture to an optical density (O.D.) of 660 nm of 0.1 in a 500 mL shake flask; (3) growing the culture at 30 degrees Celsius with agitation for 5-6 hours; (4) pouring the culture into a chilled, sealed, sterile 250 mL centrifuge bottle; Spin at 8-10K for ten minutes in Refrigerated Sorvall at 4 degrees Celsius; (5) pouring off the supernatant thoroughly and resuspending the cell pellet in an equal volume of ice-cold, sterile, deionized water; (6) centrifuging the sample again under the same conditions; (7) repeating the water wash remembering to keep everything ice-cold; (8) pouring off the supernatant thoroughly and resuspending the cell pellet in 1 mL of ice-cold, sterile 10% glycerol and transferring the cells to a chilled, sterile, 1.5 mL microcentrifuge tube; (9) spin the sample for 10 minutes in a refrigerated centrifuge; (10) pipetting off and discarding the supernatant, and resuspending the pellet in two to three times the pellet volume (200-400  $\mu$ L) of 10% glycerol; and (11) aliquoting, if necessary, the cells into chilled tubes and freezing at -70 Celsius.

pK184-KDAB, pK184-KDABH'L and pD2-KDABHL plasmid DNAs were introduced into *Corynebacterium glutamicum* host cells by the following electroporation procedure: (1) pipetting 35  $\mu$ L cell/glycerol solution onto the side wall of a chilled 0.1cm electrocuvette; (2) pipetting about 2-4  $\mu$ L of plasmid into the solution and mixing the sample by gentle pipetting up and down; (3) bringing the entire solution to the bottom of the electrocuvette by gentle tapping, avoiding the creation of bubbles; (4) keeping the sample on ice until ready for the electroschock step, wiping off any moisture on the outside of the electrocuvette prior to the electroschock administration, and shocking the cells one time at 1.5kV, 200 $\Omega$  , 25 $\mu$ F.

Cells are allowed to recover from electroporation by: (1) immediately pipetting 1 mL of warm "Recovery Media" into the electrocuvette and thoroughly mixing the solution by pipetting; (2) incubating the solution (in the electrocuvette) at 30 degrees Celsius for at least three hours for antibiotic resistance expression and cell recovery and (3) plating on selection media and incubating at 30 degrees Celsius for 3 days.

### ***Example 3***

#### ***Screening and Selection of Strains with Improved L-Lysine Production***

After 3 days of growth, single colonies of antibiotic resistant cells are individually selected to determine if there is increased L-lysine production over that which is produced by the parental host cell strain.

Recipes for all media used in these experiments are found in Tables 1 and 2. L-lysine production is determined on cultures of transformed, antibiotic resistant cells grown in shaker flasks. Briefly, seed media (Table 1), was dispensed in 20ml aliquots into deep baffled 250ml Bellco shake flasks and autoclaved for 20 minutes. After cooling to room temperature, these seed flasks were then inoculated with the strain to be tested and placed on a rotary shaker. They were incubated at 30 degrees Celsius, shaking, overnight. The following morning, the optical density (wavelength = 660nm) of each seed was recorded,



and 2ml of the culture from each seed flask was transferred to a 21 ml aliquot of FM3 media, also in a deep baffled shake flask. These "main" flasks were then returned to the shaker and incubated at 30 degrees Celsius.

After 48 hours of incubation, 1 ml of main culture was removed from each flask, and the flasks were promptly returned to the shaker. From the 1 ml sample, optical density was determined by diluting 1:50 in 0.1N HCl to dissolve the calcium carbonate present in the media. The remainder of each sample was then centrifuged to pellet cells and calcium carbonate. A 1:50 dilution of the supernatant was made in water and from this dilution the dextrose concentration was determined. Extracellular L-lysine concentrations were also determined at this time by HPLC.

High yield derivative cells may be conveniently identified by determining the percent yield from dextrose, *i.e.*, the yield of amino acid from dextrose defined by the formula  $[(\text{g amino acid produced} / \text{g dextrose consumed}) * 100] = \% \text{ yield}$ . Results are presented below in which the parental strains E12, NRRL-B11474 and ATCC 21799 are transformed with the L-lysine multi-gene isolated nucleic acid molecules of the invention identified as pK184-KDA, pK184-KDABH'L and pD(Elia)2-KDABHL. The pD2-KDABHL construct was made as in Example 4.

Strain Tested	lysine titer (g/L)	L-lysine yield (%)	Cell Deposit
NRRL-B11474	31	44	
NRRL-B11474::pK184-KDAB	32	45.7	NRRL-B-30219
NRRL-B11474::pK184-KDABH'L	36	51.8	NRRL-B-30218
NRRL-B11474::pDElia2-KDABHL	38	54.6	NRRL-B-30234
E12	1.4	0.9	
E12::pK184-KDABH'L	26.8	38	NRRL-B-30236
E12::pDElia2-KDABHL	29.8	42.5	NRRL-B-30237
ATCC21799	26.8	36.9	
ATCC21799:: pK184-KDAB	28.5	39	NRRL-B-30221
ATCC21799:: pK184-KDABH'L	31	43	NRRL-B-30220
ATCC21799:: pDElia2-KDABHL	36	50	NRRL-B-30235

Once high yield derivative cell lines are identified, the cell lines are further screened to determine that amplification of the amino acid biosynthetic pathway genes has occurred. Amplification screening may be conveniently accomplished either by (1) standard southern blot methodology to determine gene copy number or (2) by a determination of the total enzyme activity for enzymes encoded by the respective biosynthetic pathway genes of the isolated nucleic acid molecule introduced into the host cell.

A determination of gene copy number by Southern blot methodology may be done utilizing standard procedures known in the art of recombinant DNA technology, as described in the laboratory manuals referenced and incorporated herein, for example as found in J. Sambrook, E.F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

Table 1. Seed Media, SM1

Ingredient	Concentration (g/L)
Sucrose	50
Potassium Phosphate, Monobasic	0.5
Potassium Phosphate, Dibasic	1.5
Urea	3.0
Magnesium Sulfate	$5.0 \times 10^{-1}$
Polypeptone	20
Beef Extract	5.0
Biotin	$7.56 \times 10^{-4}$
Thiamine	$3.0 \times 10^{-3}$
Niacinamide	$1.25 \times 10^{-1}$
L-Methionine	$5.0 \times 10^{-1}$
L-Threonine	$2.5 \times 10^{-1}$
L-Alanine	$5.0 \times 10^{-1}$
pH	7.3

**Table 2. Main Media, FM3**

Ingredient	Concentration (g/L)
Dextrose*	60
Ammonium Sulfate	50
Potassium Phosphate, Monobasic	1.0
Magnesium Sulfate	$4.0 \times 10^{-1}$
Manganese Sulfate	$1.0 \times 10^{-2}$
Ferrous Sulfate	$1.0 \times 10^{-2}$
Biotin	$3.0 \times 10^{-4}$
Calcium Carbonate	50
Corn Steep Liquor (dissolved solids)	20
pH (adjusted with KOH)	7.4

\*Dextrose was added after autoclaving

### Example 4

#### Preparation of L-Lysine Pathway Multi-Gene Constructs

The invention further comprises additional L-lysine multi-gene constructs constructed using the PCR technique. Standard PCR and subcloning procedures were utilized, as described above, to generate 5-gene constructs similar to those in Example 1. The constructs of this example comprise the antibiotic resistance gene, chloramphenicol acyl transferase (CAT). The CAT gene was operably linked to a *Corynebacteria* phosphofructokinase promoter for expression in *Corynebacteria*.

The following steps were performed in constructing the following constructs containing the CAT gene:

1. pGEMT-*ask-asd*: ~2.6 Kb PCR product containing the *ask-asd* operon of ATCC21529 using primers *ask* and *asd* was cloned into pGEM-T (Promega pGEM-T vector systems);
2. pUC18-*ddh*: ~1.3Kb KpnI fragment of pADM21 containing *ddh* (NRRL B11474) was subcloned into pUC18 at the KpnI site;
3. pLIC1.7-*argS-lysA*: ~3Kb PCR product using template BF100

genomic DNA and primers *argS* and *lysA* was cloned into pPMG-LIC cloning vector (PharMingen);

4. pM4-*dapB-ORF2-dapA*: ~3 Kb PCR product using primers *dapB* and *dapA* was cloned into pM4 at the blunted XbaI site;

5. pFC3-*ask-asd*: ~2.6 Kb NsiI-ApaI fragment of pGEMT-*ask-asd* was cloned into pFC3 cut with PstI and ApaI;

6. pFC1-*ddh*: ~1.3 Kb SalI-EcoRI fragment of pUC18-*ddh* was cloned into pFC1 cut with SalI and EcoRI;

7. pFC1-*ddh-lysA*: ~1.5 Kb EcoRI fragment (containing the truncated *lysA* DNA) of pLIC1.7-*argS-lysA* was cloned into pFC1-*ddh* at the EcoRI site;

8. pFC1-*ddh-lysA*: ~2.1 Kb EcoRI-PstI fragment (containing the intact *lysA* DNA) of pRS6 was cloned into pFC1-*ddh* cut with EcoRI and PstI;

9. pFC5-*dapB-ORF2-dapA*: ~3.4 Kb BamHI-BglII fragment of pM4-*dapB-ORF2-dapA* was cloned into pFC5 at the BamHI site;

10. pFC5-*dapB-ORF2-dapA-ddh-lysA*: ~2.8 Kb NheI fragment of pFC1-*ddh-lysA* was cloned into pFC5-*dapB-ORF2-dapA* at the NheI site;

11. pFC5-*dapB-ORF2-dapA-ddh-lysA*: ~3.4 Kb NheI fragment of pFC1-*ddh-lysA* was cloned into pFC5-*dapB-ORF2-dapA* at the NheI site;

12. pFC3-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pFC3-KDABH'L): ~6.2 Kb NotI fragment of pFC5-*dapB-ORF2-dapA-ddh-lysA* was cloned into pFC3-*ask-asd* at the NotI site;

13. pFC3-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pFC3-KDABHL): ~6.8 Kb NotI fragment of pFC5-*dapB-ORF2-dapA-ddh-lysA* was cloned into pFC3-*ask-asd* at the NotI site;

14. pK184-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pK184-KDABH'L): ~8.8 Kb PmeI fragment of pFC3-*ask-asd-dapB-ORF2-dapA-ddh-lysA* was cloned into pK184 at the HincII or SmaI site;

15. pDElia2-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pD2-KDABHL):  
~9.4 Kb PmeI fragment of pFC3-*ask-asd-dapB-ORF2-dapA-ddh-lysA* was  
cloned into pDElia2 at the HincII site (contains the *kan* gene; is a full length 6  
gene construct);

5 16. pDElia11-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pD11-  
KDABHL): ~8.8 Kb PmeI fragment of pFC3-*ask-asd-dapB-ORF2-dapA-ddh-  
'lysA* was cloned into pDElia11 at the HincII or SmaI site (contains the *CAT*  
gene; is a truncated 6 gene construct);

10 17. pDElia11-*ask-asd-dapB-ORF2-dapA-ddh-lysA* (pD11-KDABHL):  
~9.4 Kb PmeI fragment of pFC3-*ask-asd-dapB-ORF2-dapA-ddh-lysA* was  
cloned into pDElia11 at the HincII site (contains the *CAT* gene; is a full length  
6 gene construct);

18. pDElia2: ~1.24Kb blunted PstI fragment of pUC4K ligated with  
the ~1.75Kb DraI-SspI fragment of pUC 19;

15 19. pDElia11: ~1Kb PCR product containing the chloramphenicol  
acyl-transferase gene expressed by the *C. glutamicum fda* promoter was obtained  
using primers UCdraI and UCsspI and pM4 as template and was ligated with the  
~1.75Kb DraI-SspI fragment of pUC19;

The primers utilized for the cloning procedures included:

20 *ask*: 5'-GGGTACCTCGCGAAGTAGCACCTGTCAC-3'

*asd*: 5'-GCGGATCCCCATCGCCCCTCAAAGA-3'

*dapB*: 5'-AACGGGCGGTGAAGGGCAACT-3'

*dapA*: 5'-TGAAAGACAGGGGTATCCAGA-3'

*ddh1* 5'-CCATGGTACCAAGTGC GTGGCGAG-3'

25 *ddh2* 5'-CCATGGTACCAACTGTTTCTTGC-3' Kpn I sites: GGTACC

*argS*: 5'-CTGGTTCCGGCGAGTGAGCCGACCATTCGCGAGG-3'

*lysA*: 5'-CTCGCTCCGGCGAGGTCGGAGGCAACTTCTGCGACG-3'

a primer that anneals internally to *lysA* (about 500bp upstream to the end of *lysA*).

UCdraI 5'-GGATCTTCACCTAGATCC  
UCsspl5'-CCCTGATAAATGCTTC

5 "K", "D", "A", "B," "H," "L" and "L" have the same designations as set forth above.

### Example 5

#### *Three-Fold Amplification of L-lysine Amino Acid Biosynthesis Pathway Genes*

10 For exemplary purposes only, Applicants provide herein an example wherein at least one L-lysine amino acid biosynthesis pathway gene is amplified by a factor of 3.

15 Plasmid pD11-KDABH'L (constructed in Example 4) was used in the construction of high yield derivative cell lines of the invention. For cell transformation experiments with the isolated nucleic acid molecules of the invention, the growth preparation of competent cells, and determining of relative growth may be done according to the procedure set forth above.

20 Plasmid pD11-KDABH'L DNA was introduced into NRRL-B30220 (comprising pK184-KDABH'L), using the electroporation method above. Introduction of the pD11-KDABH'L plasmid DNA into NRRL-B30220 resulted in incorporation of one copy of pD11-KDABH'L into the host cell chromosome via a single crossover homologous recombination event. The host cell comprising two copies of five genes (pD11-KDABH'L and pK184-KDABH'L) has been deposited as NRRL-B30222.

25 The amount of lysine produced by *C. glutamicum* ATCC 21799 host cells having 3 copies of 5 genes (one endogenous copy and one copy of each of pD11-KDABH'L and pK184-KDABH'L) is shown below.

09722441.112300

### L-lysine Production

Strains	L-lysine titer (g/L)	L-lysine yield (%)
ATCC 21799	26.6	45.0
NRRL-B30222	32.0	56.0

### Example 6

This example describes changing the promoter to increase the level of expression of each of these 6 genes described above. Six genes encoding six different enzymes of the biosynthetic pathway from L-aspartate to L-lysine have been inserted onto the chromosome of *Corynebacterium glutamicum*. The additional copy of each gene is from a *C. glutamicum* strain. The nucleotide sequences that regulate the level of expression (promoter) for each gene were the same as found on the *C. glutamicum* chromosome at the native loci.

Increased expression can result in increased specific activities of the enzymes and improved flux of carbon from aspartate to lysine. The yield of lysine from glucose can be improved by this technique.

The level of expression from a promoter sequence is referred to as strength. A strong promoter gives higher expression than a weak one. The mechanisms that determine the strength of a promoter have been described (Record, M.T., *et al.*, "Escherichia coli RNA Polymerase, Promoters, and the Kinetics of the Steps of Transcription Initiation," in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ASM Press (1996), pp. 792-881). Sources of promoters include nucleotide sequences from the 5' end of genes native to the *C. glutamicum* chromosome, from sequences on plasmids that replicate in *C. glutamicum*, from sequences in the genome of phage that infect *C. glutamicum*, or from sequences assembled by humans (tac, trc) and are not found in nature. Genes of ribosomal proteins, ribosomal RNAs and elongation factors show high levels of expression. The promoters of these genes are candidates for increasing expression of amino acid biosynthetic pathway genes.

Another reason for changing promoters of genes in biosynthetic pathways is to make the pathway independent of factors that control the pathway in the wild type organism. For example the native promoter of the operon that contains diaminopimelate decarboxylase of the lysine biosynthetic pathway of *C. glutamicum* can respond to arginine or lysine in the growth medium. Arginine increased transcription three-fold and lysine decreased transcription by one third (Oguiza, *et al.*, *J Bact.* 175:7356-7362 (1993)). Diaminopimelate decarboxylase activity decreased 60% in cells grown in minimal medium supplemented with 10mM lysine (Cremer *et al.*, *J Gen Microbiol.* 134:3221-3229 (1988)). Replacing the promoter of *lysA* which encodes the diaminopimelate decarboxylase is one way to make lysine biosynthesis independent of arginine and lysine levels in media.

### Example 6A

Shown below are examples of promoters that are stronger than the *askP1* promoter which regulates the gene for aspartate kinase, the first enzyme in the pathway from aspartate to lysine.

#### Beta-Galactosidase Assay of Candidate Promoters

Candidate	Specific Activity micromol/min/mg	Origin
E12	0.20	no promoter
E12/pTAC	49.80	pKK223-3
BF100	0.08	no promoter
BF100/pAD151.1	2.22	aspartokinase P1
E12	0.11	no promoter
E12/pAD151.1	1.96	aspartokinase P1
E12/5	3.46	BF100 genome
E12/7	.8.60	BF100 genome
E12/10	6.56	BF100 genome
E12/32	3.11	BF100 genome
E12/3	22.00	corynephage
E12/39	11.57	corynephage



E12/42	10.90	corynebophage
--------	-------	---------------

E12 is a *C. glutamicum* strain that does not produce lysine. E12 is a laboratory strain derived from ATCC 13059. BF100 is a high level lysine producer (NRRL-B11474). *TAC* is commercially available promoter that has been used as an example of a strong promoter. Four promoters from the *C. glutamicum* chromosome and three from a phage have been identified that are stronger than the native aspartokinase promoter.

### Example 6B

Examples of strong promoters increasing specific enzyme activity of aspartokinase when expressed in *C. glutamicum* are shown below.

Influence of IPTG on Aspartokinase activity

Strain	Regulator/promoter-gene	Inducer	nmol/min/mg
BF100	none	none	110
PD9 <trc-ask< td=""><td>lacI/trc-ask</td><td>none</td><td>103</td></trc-ask<>	lacI/trc-ask	none	103
PD9 <trc-ask< td=""><td>lacI/trc-ask</td><td>+IPTG (30 mg/L)</td><td>269</td></trc-ask<>	lacI/trc-ask	+IPTG (30 mg/L)	269
I31-2	lacI/trc-ask	none	59
I31-2	lacI/trc-ask	+IPTG (30 mg/L)	117
I31-5	lacI/trc-ask	none	59
I31-5	lacI/trc-ask	+IPTG (30 mg/L)	123
pD9 is a plasmid that replicates in <i>C. glutamicum</i> .			
I31 strains have the <i>trc-ask</i> construct integrated into the genome.			
IPTG induces genes controlled by the <i>TRC</i> promoter.			

### Example 6C

Examples of the influence of *lacI/trc-ask* on lysine production in shake flasks are shown below.

Strain	Induction	O.D.	Titre	Yield	S.P.
BF100	none	46	26	43	58
PD9 <trc-ask< td=""><td>none</td><td>49</td><td>30</td><td>49</td><td>61</td></trc-ask<>	none	49	30	49	61
PD9 <trc-ask< td=""><td>+IPTG</td><td>45</td><td>30</td><td>50</td><td>68</td></trc-ask<>	+IPTG	45	30	50	68
BF100	none	43	23	39	53
131-2	none	34	27	46	82
131-5	none	35	28	47	82
O.D. = optical density at 660nm					
Titre = grams Lysine/liter					
Yield = grams lysine made/grams dextrose consumed					
S.P. = grams lysine/O.D.					

The production of lysine by BF100 was improved by increasing the strength of the aspartokinase promoter.

### Example 7

This example demonstrates the use of vector pDElia2-ask-asd-dapA-ORF2-dapB-ddh-P1lysA (pDElia2KDABHP1L) in the construction of the high yield cell lines of the invention. The HpaI-PvaII fragment containing the P1 promoter was prepared as described in Marcel T., *et al.*, *Molecular Microbiology* 4:1819-1830 (1990). Applicants utilized standard PCR and subcloning procedures as set forth above. For cell transformation experiments with the isolated nucleic acid molecules of the invention, the growth preparation of competent cells, and determining or relative growth may be done according to the procedure set forth above.

Applicants performed the following steps in constructing the following vectors used in the L-lysine biosynthetic pathway.

1. pGEMT-*ask-asd*: ~2.6 Kb PCR product containing the *ask-asd* operon of ATCC21529 using primers *ask* and *asd* was cloned into pGEM-T (Promega pGEM-T vector systems).

2. pUC18-*ddh*: ~1.3 KpnI fragment of pADM21 containing *ddh* (BF100 locus) was subcloned into pUC18 at the KpnI site.

3. pFC3-*ask-asd*: ~2.6 Kb NsiI-ApaI fragment of pGEMT-*ask-asd* was cloned into pFC3 cut with PstI and ApaI.

4. pFC3-*dapB-ORF2-dapA*: ~2.9 Kb PCR product of NRRL-B11474 *dapB-ORF2-dapA* coding region was cloned into pFC3 at the EcoRV site.

5. pFC1-*ddh*: ~1.3 Kb PstI-EcoRI fragment of pUC18-*ddh* was cloned into pFC1 cut with PstI and EcoRI.

6. pUC19-P1: ~550 bp HpaI-PvuII fragment (containing the first promoter, P1, of the *argS-lysA* operon) of pRS6 was cloned into pUC19 at the SmaI site.

7. pUC19-P1*lysA*: ~1.45 Kb promoterless PCR product, using primer *LysA*(ATG) and *LysA*3B, of NRRL-B11474 *lysA* coding region is cloned into pUC19-P1 at the HincII site.

8. pFC1-P1*lysA*: ~2 Kb EcoRI-HindIII fragment of pUC19-P1*lysA* was cloned into pFC1 cut with EcoRI and HindIII.

9. pFC1-P1*lysA-ddh*: ~1.3 Kb EcoRI-NotI fragment of pFC1-*ddh* was cloned into pFC1-P1*lysA* cut with EcoRI and NotI.

10. pFC1-*ask-asd-ddh*-P1*lysA*: ~2.6 Kb SmaI-FseI fragment of pFC3-*ask-asd* was cloned into pFC1-*ddh*-P1*lysA* cut with SmaI and FseI.

11. pFC3-*ask-asd-dapB-ORF2-dapA-ddh*-P1*lysA* (pFC3-KDABHP1L): ~5.9 Kb SpeI fragment of pFC1-*ask-asd-ddh*-P1*lysA* was cloned into pFC3-*dapB-ORF2-dapA* at the SpeI site.

12. pDElia2-*ask-asd-dapB-ORF2-dapA-ddh*-P1*lysA* (pDElia2-KDABHP1L): ~8.8 Kb PmeI fragment of pFC3-*ask-asd-dapB-ORF2-dapA-ddh*-P1*lysA* was cloned into pDElia2 at the HincII site.

Primers used in PCR:

*lysA*(ATG): CCGGAGAAGATGTAACAATGGCTAC

*LysA3B*: CCTCGACTGCAGACCCCTAGACACC

The nucleotide sequence (SEQ ID NO:17) of the HpaI-PvuII fragment containing the promoter P1 is shown in figure 20. Results of lysine production in NRRL-B11474 comprising the pDElia2-*ask-asd-dapA-ORF2-dapB-ddh-P1lysA* (pDElia2 KDABHP1L) construct are shown below.

Strain tested	lysine titer	lysine yield (%)	cell deposit
NRRL-B11474	30	35	
NRRL-B11474::pDElia2-KDABHP1L	37	42.8	NRRL B30359

### Example 8

This example demonstrates the use of vector pDElia2<sub>FC3</sub>-*ask-asd-dapB-ddh-lysA* (pDElia2<sub>FC3</sub>KDBHL) in the construction of the high yield cell lines of the invention. The pDElia2<sub>FC3</sub>KDBHL vector comprises a truncated ORF2 gene and lacks a *dapA* gene. The ORF2 gene was cleaved at an internal ClaI site, removing the 3' region and the *dapA* gene. A promoterless *lysA* gene was obtained from NRRL-B11474. For cell transformation experiments with the isolated nucleic acid molecules of the invention, the growth preparation of competent cells, and determining of relative growth may be done according to the procedure set forth above. Applicants performed the following steps in constructing the following vectors used in the L-lysine biosynthetic pathway.

1. pGEMT-*ask-asd*: ~2.6 Kb PCR product containing the *ask-asd* operon of ATCC21529 using primers *ask* and *asd* was cloned into pGEM-T (Promega pGEM-T vector systems).

2. pFC3-*ask-asd*: ~2.6 Kb NsiI-ApaI fragment of pGEMT-*ask-asd* was cloned into pFC3 cut with PstI and ApaI.

3. pFC3-*dapB-ORF2-dapA*: ~2.9 Kb PCR product of NRRL-B11474 *dapB-ORF2-dapA* coding region was cloned into pFC3 at the EcoRV site.

4. pFC3-*dapB*: the large Clal fragment of pFC3-*dapB-ORF2-dapA* was religated.

5. pUC18-*ddh*: ~1.3 Kb KpnI fragment of pADM21 containing *ddh* (NRRL-B11474 locus) was subcloned into pUC18 at the KpnI site.

6. pFC1-*ddh*: ~1.3 Kb SalI-EcoRI fragment of pUC18-*ddh* was cloned into pFC1 cut with SalI and EcoRI.

7. pFC1-*ddh-lysA*: ~2.1 Kb EcoRI-PstI fragment (containing the intact *lysA* DNA) of pRS6 was clone into pFC1-*ddh* cut with EcoRI and PstI.

8. pFC1-*ask-asd-ddh-lysA*: ~2.6 Kb SwaI-FseI fragment of pFC3-*ask-asd* was cloned into pFC1-*ddh-lysA* cut with SwaI and FseI.

9. pFC3-*ask-asd-dapB-ddh-lysA*: ~6 Kb SpeI fragment of pFC1-*ask-asd-ddh-lysA* was cloned into pFC3-*dapB* at the SpeI site.

10. pDElia2<sub>FC5</sub>-*ask-asd-dapB-ddh-lysA* (pDElia2<sub>FC5</sub>-KDBHL): ~7.3 Kb NotI-PmeI fragment of pFC3-*ask-asd-dapB-ddh-lysA* was cloned into pDElia2<sub>FC5</sub> cut with NotI and PmeI.

11. pDElia2<sub>FC5</sub>: the small PvuII fragment of pFC5 was ligated with the large PvuII fragment of pDElia2.

Results of lysine production in NRRL-B11474 comprising the pDElia2<sub>FC5</sub>-*ask-asd-dapB-ddh-lysA* (pDElia2<sub>FC5</sub>KDBHL) are shown below.

Strain tested	lysine titer	lysine yield (%)	cell deposit
NRRL-B11474	31	49	
NRRL-B11474::pDElia2 <sub>PC5</sub> -KDBHL	37.8	58	NRRL B30360

\* \* \* \* \*

5            Having now fully described the present invention in some detail by way  
of illustration and example for purposes of clarity of understanding, it will be  
obvious to one of ordinary skill in the art that same can be performed by  
modifying or changing the invention with a wide and equivalent range of  
10 conditions, formulations and other parameters thereof, and that such  
modifications or changes are intended to be encompassed within the scope of the  
appended claims.

15            All publications, patents and patent applications mentioned in this  
specification are indicative of the level of skill of those skilled in the art to which  
this invention pertains, and are herein incorporated by reference to the same  
extent as if each individual publication, patent or patent application was  
specifically and individually indicated to be incorporated by reference.

00720441 720000

***What Is Claimed Is:***

1. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2.
2. An isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide sequence of claim 1.
3. The isolated polynucleotide molecule of claim 2 comprising a nucleic acid having the sequence of SEQ ID NO:1.
4. A vector comprising the isolated polynucleotide molecule of claim 2.
5. A host cell comprising the vector of claim 4.
6. A method comprising:
  - (a) transforming a *Corynebacterium* species host cell with the polynucleotide molecule of claim 2, wherein said isolated polynucleotide molecule is integrated into said host cell's chromosome thereby increasing the total number of said amino acid biosynthetic pathway genes in said host cell chromosome, and
  - (b) selecting a transformed host cell.
7. The method of claim 6 further comprising screening for increased amino acid production.
8. The method of claim 6 wherein said polynucleotide molecule further comprises at least one of the following:
  - (a) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *asd* amino acid sequence;

(b) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapA* amino acid sequence;

(c) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *dapB* amino acid sequence;

5 (d) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ddh* amino acid sequence; and

(e) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence;

10 (f) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *lysA* amino acid sequence; and

(g) a nucleic acid molecule encoding a *Corynebacterium species* lysine pathway *ORF2* amino acid sequence.

9. The method of claim 8 further comprising screening for increased amino acid production.

15 10. The method of claim 6, wherein said isolated polynucleotide molecule further comprises at least one of the following:

(a) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

20 (b) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

(c) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

(d) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10;

25 (e) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21;

(f) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14;

00722441 12600



(g) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

11. The method of claim 6, wherein said isolated polynucleotide molecule further comprises the following:

5 (a) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

(b) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

10 (c) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8; and

(d) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

12. The method of claim 6, wherein said isolated polynucleotide molecule further comprises the following:

15 (a) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

(b) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

20 (c) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

(d) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; and

(e) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

25 13. The method of claim 6, wherein said isolated polynucleotide molecule further comprises the following:

(a) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

(b) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

(c) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

5 (d) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10;

(e) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; and

10 (f) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

14. The method of claim 6, wherein said isolated polynucleotide molecule further comprises the following:

(a) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

15 (b) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

(c) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

20 (d) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10;

(e) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and

(f) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

25 15. The method of claim 6 further comprising:

(a) growing said transformed host cell in a medium; and

(b) purifying an amino acid produced by said transformed host cell.

09722441-112600

16. An isolated polynucleotide molecule comprising:

- (a) the polynucleotide molecule of claim 2; and
- (b) at least one additional *Corynebacterium* species lysine pathway

gene selected from the group consisting of:

- (i) a nucleic acid molecule encoding an *asd* polypeptide;
- (ii) a nucleic acid molecule encoding a *dapA* polypeptide;
- (iii) a nucleic acid molecule encoding a *dapB* polypeptide;
- (iv) a nucleic acid molecule encoding a *ddh* polypeptide;
- (v) a nucleic acid molecule encoding a *lysA* polypeptide;
- (vi) a nucleic acid molecule encoding a *lysA* polypeptide; and
- (vii) a nucleic acid molecule encoding an *ORF2* polypeptide.

17. The isolated nucleic acid molecule of claim 16, wherein:

- (a) said *asd* polypeptide is SEQ ID NO:4;
- (b) said *dapA* polypeptide is SEQ ID NO:6;
- (c) said *dapB* polypeptide is SEQ ID NO:8;
- (d) said *ddh* polypeptide is SEQ ID NO:10;
- (e) said *lysA* polypeptide is SEQ ID NO:21;
- (f) said *lysA* polypeptide is SEQ ID NO:14; and
- (g) said *ORF2* polypeptide is SEQ ID NO:16.

18. An isolated polynucleotide molecule comprising:

- (a) the polynucleotide molecule of claim 2;
- (b) a nucleic acid molecule encoding the *asd* amino acid sequence of

SEQ ID NO:4;

- (c) a nucleic acid molecule encoding the *dapA* amino acid sequence

of SEQ ID NO:6;

- (d) a nucleic acid molecule encoding the *dapB* amino acid sequence

of SEQ ID NO:8; and

- (e) a nucleic acid molecule encoding the *ORF2* amino acid sequence

of SEQ ID NO:16.

19. An isolated polynucleotide molecule comprising:

- (a) the polynucleotide molecule of claim 2;
- (b) a nucleic acid molecule encoding the *asd* amino acid sequence of

SEQ ID NO:4;

5 (c) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

(d) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

10 (e) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10; and

(f) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

20. An isolated polynucleotide molecule comprising:

- (a) the polynucleotide molecule of claim 2;

15 (b) a nucleic acid molecule encoding the *asd* amino acid sequence of SEQ ID NO:4;

(c) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

20 (d) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

(e) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10;

(f) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:21; and

25 (g) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

09722441.12800

21. An isolated polynucleotide molecule comprising:

- (a) the polynucleotide molecule of claim 2;
- (b) a nucleic acid molecule encoding the *asd* amino acid sequence of

SEQ ID NO:4;

(c) a nucleic acid molecule encoding the *dapA* amino acid sequence of SEQ ID NO:6;

(d) a nucleic acid molecule encoding the *dapB* amino acid sequence of SEQ ID NO:8;

(e) a nucleic acid molecule encoding the *ddh* amino acid sequence of SEQ ID NO:10;

(f) a nucleic acid molecule encoding the *lysA* amino acid sequence of SEQ ID NO:14; and

(g) a nucleic acid molecule encoding the *ORF2* amino acid sequence of SEQ ID NO:16.

22. The isolated polynucleotide molecule of claim 18 comprising pK184-KDAB.

23. The isolated polynucleotide molecule of claim 20 comprising pD11-KDABH'L.

24. The isolated polynucleotide molecule of claim 21 comprising pD2-KDABHL.

25. A vector comprising the polynucleotide molecule of claim 16.

26. A host cell comprising the vector of claim 25.

27. The host cell of claim 26 wherein said host cell is a *Brevibacterium* selected from the group consisting of *Brevibacterium flavum* NRRL-B30218, *Brevibacterium flavum* NRRL-B30219, *Brevibacterium lactofermentum*

NRRL-B30220, *Brevibacterium lactofermentum* NRRL-B30221, *Brevibacterium lactofermentum* NRRL-B30222, *Brevibacterium flavum* NRRL-30234 and *Brevibacterium lactofermentum* NRRL-30235.

28. The host cell of claim 26 wherein said host cell is *Escherichia coli* DH5  $\alpha$  MCR NRRL-B30228.

29. The host cell of claim 26 wherein said host cell is a *C. glutamicum* selected from the group consisting of *C. glutamicum* NRRL-B30236 and *C. glutamicum* NRRL-B30237.

30. A method of producing lysine comprising culturing the host cells of claim 5 wherein said host cells comprise one or more of:

(a) increased enzyme activity of one or more lysine biosynthetic pathway enzymes compared to the genetically unaltered host cell;

(b) one or more copies of each gene encoding a lysine biosynthetic pathway enzyme; and,

(c) alteration of one or more transcription factors regulating transcription of one or more genes encoding a lysine biosynthetic pathway enzyme, wherein said host cell produces lysine in said culture medium.

31. The method of claim 30 wherein said increased enzyme activity comprises overexpressing one or more genes encoding one or more lysine biosynthetic pathway enzymes.

32. The method of claim 31 wherein said one or more genes are operably linked directly or indirectly to one or more promoter sequences.

33. The method of claim 32 wherein said operably linked promoter sequences are heterologous, endogenous, or hybrid.

34. The method of claim 33 wherein said promoter sequences are one or more of: a promoter sequence from the 5' end of genes endogenous to *C. glutamicum*, a promoter sequence from plasmids that replicate in *C. glutamicum*, and, a promoter sequence from the genome of phage which infect *C. glutamicum*.

5 35. The method of claim 33 or 34 wherein one or more of said promoter sequences are modified.

36. The method of claim 35 wherein said modification comprises truncation at the 5' end, truncation at the 3' end, non-terminal insertion of one or more nucleotides, non-terminal deletion of one or more nucleotides, addition of one or more nucleotides at the 5' end, addition of one or more nucleotides at the 3' end, and, combinations thereof.

37. The method of claim 30 wherein said increased enzyme activity results from the activity of one or more modified lysine biosynthetic pathway enzymes wherein said enzyme modification results in a change in kinetic parameters, allosteric regulation, or both, compared to the enzyme lacking said modification.

38. The method of claim 37 wherein said change in kinetic parameters is a change in  $K_m$ ,  $V_{max}$  or both.

39. The method of claim 37 wherein said change in allosteric regulation is a change in one or more enzyme allosteric regulatory sites.

40. The method of claim 37 wherein said modification is a result of a change in the nucleotide sequence encoding said enzyme.

41. The method of claim 40 wherein said change in said nucleotide sequence is an addition, insertion, deletion, substitution, or a combination thereof, of one or more nucleotides.

42. The method of claim 30 wherein said alteration of one or more transcription factors comprises one or more mutations in transcription inhibitor proteins, one or more mutations in transcription activator proteins, or both, wherein said one or more mutations increases transcription of the target nucleotide sequence compared to the transcription by said one or more transcription factors lacking said alteration.

43. The method of claim 42 wherein said one or more mutations is a change in said nucleotide sequence encoding said transcription factor.

44. The method of claim 43 wherein said change in said nucleotide sequence is an addition, insertion, deletion, substitution, or a combination thereof, of one or more nucleotide.

45. An isolated polypeptide, wherein said polypeptide comprises an amino acid sequence having at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 19.

46. The polypeptide of claim 45, wherein said polypeptide has the amino acid sequence of SEQ ID NO: 19.

47. An isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide of claim 45.

48. The isolated polynucleotide molecule of claim 47 comprising a nucleic acid having the sequence of SEQ ID NO: 18.

49. A vector comprising the polynucleotide molecule of claim 47.

50. A host cell comprising the vector of claim 49.

51. The host cell of claim 50 wherein said host cell is NRRL B30360.



52. A method comprising:

(a) transforming a *Corynebacterium* species host cell with the polynucleotide molecule of claim 47, and

(b) selecting a transformed host cell.

5 53. An isolated polypeptide wherein said polypeptide comprises a polypeptide having at least 95% sequence identity to the amino acid sequence of SEQ ID NO:21.

54. The polypeptide of claim 53 having the amino acid sequence of SEQ ID NO:21.

10 55. An isolated polynucleotide molecule comprising a nucleotide sequence encoding the polypeptide of claim 53.

56. The isolated polynucleotide molecule of claim 55 comprising a nucleic acid having the sequence of SEQ ID NO: 20.

57. A vector comprising the polynucleotide molecule of claim 55.

15 58. A host cell comprising the vector of claim 57.

59. The host cell of claim 58 wherein said host cell is selected from the group consisting of NRRL B30218, NRRL B30220 and NRRL B30222.

60. A method comprising:

(a) transforming a *Corynebacterium* species host cell with the polynucleotide molecule of claim 55, and

(b) selecting a transformed host cell.

09722441.112800

5

62. The polynucleotide of claim 61 where said promoter sequence has the nucleotide sequence of SEQ ID NO: 17.

63. The isolated polynucleotide molecule of claim 61 wherein said promoter is operably directly linked to the LysA gene.

64. A vector comprising the isolated polynucleotide of claim 61.

65. A host cell comprising the vector of claim 64.

10

66. The host cell of claim 65 wherein said host cell is NRRL B30359.

67. A method comprising:

(a) transforming a *Corynebacterium* species host cell with the polynucleotide molecule of claim 61, and

(b) selecting a transformed host cell.

## Increased Lysine Production by Gene Amplification

### *Abstract*

5 The invention provides methods to increase the production of an amino acid from *Corynebacterium* species by way of the amplification of amino acid biosynthetic pathway genes in a host cell chromosome. In a preferred embodiment, the invention provides methods to increase the production of L-lysine in *Corynebacterium glutamicum* by way of the amplification of L-lysine biosynthetic pathway genes in a host cell chromosome. The invention also provides novel processes for the production of an amino acid by way of the amplification of amino acid biosynthetic pathway genes in a host cell chromosome and/or by increasing promoter strength. In a preferred embodiment, the invention provides processes to increase the production of L-lysine in *Corynebacterium glutamicum* by way of the amplification of L-lysine biosynthetic pathway genes in a host cell chromosome. The invention also provides novel isolated nucleic acid molecules for L-lysine biosynthetic pathway genes of *Corynebacterium glutamicum*.

```

graph TD
    LAsp[L-Aspartate] -->|Aspartate kinase (Ask)| LAspP[L-Aspartylphosphate]
    LAspP -->|Aspartatesemialdehyde dehydrogenase (asd)| LAspS[L-Aspartatesemialdehyde]
    LAspS -.->|L-Methionine synthase (dapA)| LThr[L-Threonine]
    LAspS -->|Dihydropicolinate reductase (dapB)| L23D[L-2,3-Dihydropicolinate]
    L23D -->|Dihydropicolinate succinylase (dapD)| LPD[L-Piperidine-2,6-dicarboxylate]
    LPD -->|N-Succinyl-amino-Ketopimelate transaminase (dapC)| NS2AK[N-Succinyl-2-amino-6ketopimelate]
    NS2AK -->|N-Succinyl-diamino-pimelate desuccinylase (dapE)| NS26L[L,N-Succinyl-2,6-L,L-diaminopimelate]
    NS26L -->|Diaminopimelate epimerase (dapF)| LLDP[L,L-Diaminopimelate]
    LLDP -->|Diaminopimelate decarboxylase (lysA)| DLD[L,D-Diaminopimelate]
    DLD -->|Permease| LLys[L-Lysine]
    LLys -->|L-Lysine decarboxylase (lysB)| LLys2[L-Lysine]
    LAsp -.-> LThr
    LAsp -.-> LIs[L-Isoleucine]
  
```

Metabolic pathway of L-Aspartate in *E. coli*:

- L-Aspartate is converted to L-Aspartylphosphate by Aspartate kinase (Ask).
- L-Aspartylphosphate is converted to L-Aspartatesemialdehyde by Aspartatesemialdehyde dehydrogenase (asd).
- L-Aspartatesemialdehyde can be converted to L-Threonine by L-Methionine synthase (dapA) or to L-2,3-Dihydropicolinate by Dihydropicolinate reductase (dapB).
- L-2,3-Dihydropicolinate is converted to L-Piperidine-2,6-dicarboxylate by Dihydropicolinate succinylase (dapD).
- L-Piperidine-2,6-dicarboxylate is converted to N-Succinyl-2-amino-6ketopimelate by N-Succinyl-amino-Ketopimelate transaminase (dapC).
- N-Succinyl-2-amino-6ketopimelate is converted to N-Succinyl-2,6-L,L-diaminopimelate by N-Succinyl-diamino-pimelate desuccinylase (dapE).
- N-Succinyl-2,6-L,L-diaminopimelate is converted to L,L-Diaminopimelate by Diaminopimelate epimerase (dapF).
- L,L-Diaminopimelate is converted to D,L-Diaminopimelate by Diaminopimelate decarboxylase (lysA).
- D,L-Diaminopimelate is converted to L-Lysine by Permease.
- L-Lysine is then converted to L-Lysine by L-Lysine decarboxylase (lysB).

Figure 1

Nucleotide

sequence of ATCC21529 ask (SEQ ID NO:1)

```

1  GTGGCCCTGC TCGTACAGAA ATATGCCGGT TCTCGCTTG AGACTGCCGA
51  ACGCATTTAGA AACGTCGCTG AACGGAATCGT TCCACCAAG AAGGCTGGAA
101 ATCATGTCGT GGTGTCTGCG TCCGCAATGG GAGACCCAC GGATGAACCT
151 CTAGAACTTG CAGCGGCAGT GAATCCCGTT CCGCCAGCTC GTGAATGGA
201 TATGCTCCTG ACTGCTGGTG AGCGTATTTC TAACGCTCTC GTCGCCATGG
251 CTATTGAGTC CCTTGGCGCA GAAGCTCAAT CTTTCACTGG CTCTCAGGCT
301 GGTGTGCTCA CCACCGAGCG CCACGGAAAC GCAOGCATTG TTGACGTAC
351 ACCGGGTCGT GTGCGTGAAG CACTCGATGA GGGCAAGATC TGCATTGTTG
401 CTGGTTTTCA GGGTGTTAAT AAGAAACCC GCGATGTCAC CACGTTGGGT
451 CGTGGTGGTT CTGACCCAC TGCAGTTGCG TTGGCAGCTG CTTTGAACGC
501 TGATGTGTGT GAGATTTACT CGGACGTTGA CGGTGTGTAT ACCGCTGACC
551 CGCGCATCGT TCCTAATGCA CAGAAGCTGG AAAAGCTCAG CTTGGAAGAA
601 ATGCTGGAAC TTGCTGCTGT TGCTCCAAAG ATTTGGTGC TGCGAGTGT
651 TGAATACGCT CGTGCTTCA ATGTGCCACT TCGGTACGC TCGTCTTATA
701 GTAATGATCC CGGCACTTTG ATTGCGGCT CTATGGAGGA TATTCCTGTG
751 GAAGAAGCAG TCCTTACCGG TGTGCCAACC GACAAGTCCG AAGCCAAAGT
801 AACCGTTCTG GGTATTTCCG ATAAGCCAGG CGAGGCTGCC AAGGTTTTCC
851 GTGCGTTGGC TGATGCAGAA ATCAACATTG ACATGGTTC TCGAAGCGTC
901 TCCTCTGTGG AAGACGGCAC CACGACATC ACCTTCACCT GCCCTGCGC
951 TGACGGAGCG CGTGCGATGG AGATCTTGAA GAAGCTTCAG GTTCAGGGCA
1001 ACTGGACCAA TGTGCTTTAC GACGACCAGG TCGGCAAAGT CTCCTCGTG
1051 GGTGCTGGCA TGAAGTCTCA CCCAGGTGTT ACCGCAGAGT TCATGGAAGC
1101 TCTGCGGAT GTCAACGTGA ACATCGAATT GATTTCATC TCTGAGATCC
1151 GCATTTCCGT GCTGATCCGT GAAGATGATC TGGATGCTGC TGCAGTGA
1201 TTGCATGAGC AGTTCAGCT GGGCGCGGAA GACGAAGCCG TCGTTATGC
1251 AGGCACCGGA CGCTAA

```

Fig. 2

0972244.1.12300

SEQ ID NO: 2

pro amino acid sequence of ATCC 21529 act.

1 GTGGCCCTGTCGTACAGAAATATGGCGTCTTCGCTTGAGAGTGGGAAGCAATTAGA 60  
K A L V V O K Y G G S S L E S A E P : R  
61 ACGTCGCTGAACGGATCGTTGCCACCAAGAGGCTGAAATGATGCTGGTCTCTGCG  
N V A E R I V A T K K A G N D V V V V C  
121 TCCGCAATGGGAGACACCAACGGATGAACCTCTAGAACTTGACCGGCACTGAATCCCGTT  
S A M G D T T D E L L E L A A A V N P V  
CCGCCAGCTCGTGAAATGGATATGCTCTGACTGCTGGTGAGCGTATTTCTAACGCTCTC  
181 P P A R E M D M L L T A G E R I S N A L  
GTGCGCATGGCTATTGAGTCCCTTGGCGGAGAGGCTCAATCTTCACTGGCTCTCAGGCT  
241 V A M A I E S L G A E A Q S F T G S Q A  
GGTGTGCTCAACACCGAGCGCCACGGAACGACGCAATGTTGACGTCACACCGGGTCGT  
301 G V L T T E R H G N A R I V D V T P G R  
GTGCGTGAAGCACTCGATGAGGGCAAGATCTGCATYCTGCTGGTTTTCAGGGTGTAAAT  
361 V R E A L D E G K I C I V A G F Q G V N  
NAAGAAACCCGCGATGTCACCAAGTGGGCTGGTGGTCTGACACCACTGCACTGGCG  
421 K E T R D V T T L G R G G S D T T A V A  
TTGGCAGCTGCTTTGAACGCTGATGTGTGTGAGATTACTCGGACGTTGAGGGTGTGTAT  
481 L A A A L N A D V C E I Y S D V D G V Y  
ACCGCTGACCCGCGCATCGTTCTTAATGCACAGAGCTGGAAGAGCTCAGCTTCAAGAA  
541 T A D P R I V P N A Q K L E K L S F E E  
ATGCTGGAACTGTGCTGCTGTGGCTCCAGATTTTGGTCTGCGCAGTGTGAATACGCT  
601 M L E L A A V G S K I L V L R S V E Y A  
CGTGCACTCAATGTGCCACTTGGCTACGCTCGTCTTATAGTAATGATCCCGGCACTTTC  
661 R A F N V P L R V R S S Y S N D P G T L  
ATTGCCGGCTCTATGGAGGATATTCCTGTGGAAGACAGTCCCTACCGGTGTCCCAACC  
721 I A G S M E D I P V E E A V L T G V A T  
GACAAGTCCGAAGCCAAAGTAACGCTTCTGGGTATTTCCGATAAGCCAGGCGAGGCTGCC  
781 D K S E A K V T V L G I S D K P G E A A

Fig 3A

390x  
114

0572441-115610

Seq ID NO:2  
(cont)

```

841 AAGGTTTTCCGTGCGTTGGCTGATGCAAAATCAACATTGACATGGTTCTTCAGAACGTC----- 900
    K V F R A L A D A E I I D M V L E N Y
901 TCCTCTGTGGAGACGGCACCACCGACATCACGTTACCTGCCCTCGCGCTTACGGACGGC----- 960
    S S V E D G T T D E I F T C P R A D G R
961 CGTGCGATGGAGATCTTGAAGAAGCTTCAGGTTCAAGGCAACTGGACCAATGTGCTTTAC----- 1020
    R A M E I L K K L Q V Q G N W T N V L Y
1021 GACGACCAGGTCGGCAAAGTCTCCCTCGTGGTGCTGGCATGAAGTCTCACCCAGGTGTT----- 1080
    D D Q V G K V S L V G A G M K S H P G V
1081 ACCGCAGAGTTCATGGAAGCTCTGCGCGATGTCAACSTGAACATCGAATTGATTCCATC----- 1140
    T A E F M E A L R E V N V N I E L I S I
1141 TCTGAGATCCGCATTTCCGTGCTGATCCSTGAAGATGATCTGGATGCTGCTSCACGTGCA----- 1200
    S E I R I S V L I R E D D L D A A A R A
1201 TTGCATGAGCAGTTCCAGCTGGGCGGCGAGACGAAGCCGTCTGTTATGCAGGCACCGGA----- 1260
    L H E Q F Q L G G E D E A V V Y A S T G
1261 CGCTAA----- 1266
    R *

```

Fig. 38

0972444.112600

3

Nucleotide.                      sequence of ATCC21529 as<sup>2</sup> (SEQ ID NO:2)

```

1  ATGACCACCA TCGCAGTTGT TGETGCAACC GGCAGGTCG GCCAGGTTAT
51  CGGCACCTTT TTGCAAGAGC GCAATTCCCG AGCTGACACT GTTCGTTTCT
101 TTGCTTCCCG GCGTTCGCGA GGCCTAAGA TTGAATCCG TGGCACGGAA
151 ATCGAGGTAG AAGACATTAC TCAGCAACC CAGGAGTCCG TCAAGGGCAT
201 CSACGTTGCG TTGTTCTCTG CTGGAGGCAC CGCTTCCAAG CAGTACGCTC
251 CACTGTTTGC TGCTGCAGGC GCGACTGTTG TGGATAACTC TTCTGCTTGG
301 CGCAAGGACG ACGAGGTTCC ACTAATCGTC TCTGAGGTGA ACCCTTCCGA
351 CAAGGATTCC CTGGTCAAGG GCAATTATGC GAATCCTAAC TGCACCACCA
401 TGGCTGCAAT GCCAGTGGTG AAGCCACTGC AGGATGCCGC TGGTCTTGTA
451 AAGCTTCACG TTCTCTTTA CAGGCTGTT TCCGGTTCG GTCTTGCAAG
501 TGTGGAAACC TTGGCAAGC AGGTGCTGC AGTTGGGAC CACARCGTTG
551 AGTTCGTCCA TGATGGACAG GCTGCTGACG CAGGCGATGT CGGACCTTAC
601 GTTTCCCCAA TCGCTTACAA CGTCTTGCCA TCGCCGGGAA ACCTCGTCGN
651 TGACGGCACC TTCGAAACCG ACGAAGACCA GAGCTCGCG AACGAATCCC
701 GCAAGATTCT CGGCCCTCCA GACCTCAAGG TCTCAGGCAC CTGGGTCGCG
751 GTGCCGGTTT TCACCGGCCA CACGCTGACC ATTCAAGCCG AATTGACAA
801 GGCAATCACC GTCAGCAGG GCGAGGAGAT CTGGGTGCC GCTTCAGCGG
851 TCGAGCTTGT CGACGTCCCA ACCCCACTTG CAGCTGCCGG CATTCAGCAA
901 TCCCTCGTTG GACGCATCCG TCAGSACTCC ACTGTCGACG ACAACCGCGG
951 TCTGTTCTC GTCGTATCTG GCGATACCTC TCGCAAGGGC GCAGCACTGA
1001 ACACCATTCA GATTGCTGAG CTGCTGGTTA AGTAA

```

Fig. 4

09722441.1:2800



SEQ. ID. No: 4 amino acid sequence of pTCC 21529 asd

```

1  ATGACCACCATCCGAGTGTGTGTCGACCGGCGAGGTCCGGCAGGTTATGCGCACTTT 60
   N T T I A V V G A T G Q V G Q V M R T F
61  TTGGAGAGCGCAATTCCGACGTGACGTGTGTGTTCTTTGCTCCCGCGTTCGCGA 120
   L E E R N F P A D T V R F F A S P R S A
121  GGCGGTAGATTGAATTCGCTGGCAGGAAATCGAGGTAGAAGACATTACTCAGGCAAC 180
   G R K I E P A C T E I E V E D I T O A T
181  GAGGAGTCCCTCAAGGGCATCAGCTTTCGCTGCTCTGTGAGGACACCGCTTCCAAG 240
   E E S L K C I D V A L F S A G G T A S K
241  CAGTACGCTCCACTGTTTGTCTGCTGAGGCGTCACTGTTGTGGATAACTCTTCTGCTGG 300
   Q Y A E L F A A A G A T V V D N S S A W
301  CGCAGGACGACGAGGCTTCGCTAATGCTCTGTGAGGTGAACCTTCGACAGAGGATCC 360
   R K D D E V P L I V S E V N F S D X D S
361  CTGGTCAGGGCATTATTGCGAATGCTAACTGCACCCAGATGCTGCAATCGAGTGCTG 420
   L V K C I I A N P H C T T H A A M P V L
421  AAGCCACTGCGAGATGCGCGTGTCTGTAAAGCTTCAAGTTTCTCTTACCAAGCTGTT 480
   K P L R D A A G L V L L H V S S T Q A V
481  TCCCGTTCTGGTCTTCAGAGTGTGCAACCTTGGCAAGCAGGTGTCTGCACTTGGCGAC 540
   S G S G L A C V E T L A K Q V A A V G D
541  CACAAGCTTGAGTTGCTGCATGATGCAAGTCTGCTGACGCAAGGATATCGGACCTTAC 600
   H N V E F V H O C Q A A D A G D V G R Y
601  GTTCCCGCAATGCTTACAAGTGTGCTGCCATTCGCGCGAAGCTGCTGATGACGGGACC 660
   V S P I A Y N V L P F A C H L V D U G T
661  TTCGAAGCGACGAGAGCAGAAAGCTGGGCAAGCAATCCCGCAGATTCTGGGCTCCCA 720
   F E T D E E Q K L A K E S R K I L G L P
721  GAGCTCAAGTCTCAGGCACCTTCGCTCCGCTGCCGTTTTCACCGCGCAGACCTGACC 780
   D L K V S G T C V R V P V E T C H Y L T
781  ATTCAAGCGCAATTGCGACAGCAATCAGCTGCGACGCGCAGGAGATGTTGGTGCC 840
   I H A E F D K A I T V E O A Q E I L G A
841  GCTTACGGGCTCAGCTTTCGCAAGTCCCAAGCCCACTTCAGCTCCGCGCAATGACGAA 900
   A S G V E L V D V P T P L A A A C I O R

```

Fig 5A

09722441.112800

SEQ ID NO:4  
(Cont)

TCCCTCGTTGGA<sup>6</sup>GCATCCGTCAGGACTCCACTGTCCGACGACAACCGCGGTCTGGTTCTC  
 901 -----+-----+-----+-----+-----+ 960  
 S L V G R I R Q D S T V D D N E G L V L  
 GTCGTATCTGGCGATAACCTTCGCAAGGGCGCAGCACTGAACACCATTCAGATTGCTGAG  
 961 -----+-----+-----+-----+-----+ 1020  
 V V S G D N L R K G A A L N T I Q I A E  
 CTGCTGGTTAAGTAA  
 1021 -----+----- 1035  
 L L V K \*

Fig 5B

008211-1442260

Nucleotide, sequence of dapa (SEQ ID NO:5)

```

1 ATGAGCACAG GTTTAAGAGC TAAGACCGGA GTAGAGCACT TCGGCACCGT
51 TGGAGTAGGA ATGGTTACTC CATTACCGGA ATCCGAGAGC ATCGATATCG
101 CTGCTGGCCG CGAAGTCGGG GCTTATTTCG TTGATAAGGG CTGGATTCTT
151 TTGGTTCTCG CGGGCACGAC TGGTGATCC CCAACGACAA CCGCCGCTGA
201 AAACACTAGAA CTGCTCAAGG CCGTTCGTGA GGAATTGGG GATCGGCGGA
251 AGCTCATCGC CGCTCTGGGA ACCAACAAAC CCGGACATC TGTGAACTT
301 GCGGAAGCTG CTGCTTCTGC TGGCGCAGAC GGCCTTTTGG TTCTAATCTC
351 TTATTACTCC AAGCCGAGCC AAGAGGGATT GCTGGGCGAC TTCGGTGCAA
401 TTGCTGCAGC AACAGAGGTT CCAATTGTTC TCTATGACAT TCCTGGTCGG
451 TCAGGTATTC CAATTGAATC TGATACCATG AGACGCTGA GTGAATTACC
501 TACGATTTTG CGGTCGAAGG ACGCCAAAGG TGACCTCGTT CGAGCCACCT
551 CATTGATCAA AGAAACGGGA CTTGCCCTGT ATTCAAGCGA TGACCCACTA
601 AACCTTGTTT GCGTTGCTTT GGGCGGATCA GGTTCATTT CCGTAATTGG
651 ACATGCAGCC CCCACAGCAT TACGTGAGTT GTACACAGC TTCGAGGAAG
701 GCGACCTCGT CCGTGCAGG GAAATCAACG CCAAACTATC ACCGCTGTGA
751 GCTGCCAAG GTCCTTGGG TGGAGTCAGG TTGGCAAAAG CTGCTCTGG
801 TCTGCAGGGC ATCAACGTAG GAGATCCTCG ACTTCAAATT ATGGCTCCAA
851 ATGAGCAGGA ACTTGAGGCT CTCGAGAGAG ACATGAAAA AGCTGGAGTT
901 CTATAA

```

Fig 6

09722441.112800

aa sequence  
of dapft

Fig 7

**DOBET-F-460**

Nucleotide sequence of *dapB* (SEQ ID NO: 7)

```

1  ATGGGAATCA AGGTTGGCCT TCTCGGAGCC AAGGCGCGTG TTGGTCAAAC
51  TATTGTGGCA GCAATCAATG AGTCCGACGA TCTGGAGCTT GTTGCAAGA
101 TCGGCGTCTGA CGATGATTTG AGCCTTCTGG TAGACAACGG CGCTGAAGTT
151 CTCGTTGACT TCACCACTCC TAACGCTGTG ATGGGCAACC TGGAGTTCTG
201 CATCAACAAC GGCATTTCCT CGGTGTGTGG AACCACGGGC TTCGATGATG
251 CTGCTTTGGA GCAAGTTTCG GcCTGGCTTG AAGGAAAGA CAATGTGCGT
301 GTTCTGATCG CACCTTACTT TGCTATCTCT GCGGTGTTGA CCAATGTCCT
351 TTCCAAGCAG GCTGCCCGCT TCTTGAATC AGCTGAAGTT ATTGAGCTGC
401 ACCACCCCAJ CAAGCTGGAT GCACCTTCAG GCACCGCGAT CCACACTGCT
451 CAGGGCATTG CTGCGGCAGC CAAAGPAGCA GGCATGGACG CACAGCCAGA
501 TGGACCCGAG CAGGCACCTG AGGTTCCCGG TGGCCCAAGC GTAGATGGAA
551 TCCCAATTCA CCGAGTCGCG ATGTCCGGCA TGTTGCTCA CGAGCAAGTT
601 ATCTTTGGCA CCCAGGGTCA GACCTTGACC ATCAAGCAGG ACTCCTATGA
651 TCGCAACTCA TTTCACACAG GTGTCTTGST GGTGTGCGC AACATTGCAC
701 AGCACCCAGG CCTAGTCGTA GGACTTGAGC ATTACCTAGG CCTGTAA

```

Fig 8

0972441.1.12800

SRQID NO. 8 Amino acid sequence of dapB

1 ATGGGAATCAAGTTGGCGTTCTCGGAGCCAAAGCCCGTGTGGTCAAACACTATTGTGGCA 60  
 M G I K V G V L G A K G R V G Q T I V A  
 61 GCAGTCAATGAGTCCGACGATCTGGAGCTTGTTCAGAGATCGCGCTCGACGATGATTTC 120  
 A V N E S D D L E L V A E I G V D D D L  
 121 AGCCTTCTGGTAGACAACGGCGCTGAAGTGTGCTTGACTTCACCACTCCTAACGCTGTG 180  
 S L L V D N G A E V V V D F T T P N A V  
 181 ATGGGCAACCTGGAGTTCTGCATCAACAACGGCAITTCGCGGTGTGGGAACACGGGC 240  
 M G N L E F C I N N G I S A V V G T T G  
 241 TTCGATaATGCTCGTTTGGAGCAGGTTGCGGcCTGGCTGAAGGAAGAACAATGTGCGT 300  
 F D N A R L E Q V R A W L E G K D N V G  
 301 GTTCTGATCGCACCTAACTTTGCTATCTCTGCGGTGTGACCATGGTCTTTTCGAAGCAG 360  
 V L I A P N F A I S A V L T M V F S K Q  
 361 GCTGCCCGCTTCTCGAATCAGCTGAAGTTATTGAGCTGCACCAACCCCAACGCTGGAT 420  
 A A R E F E S A E V I E L H H P N K L D  
 421 GCACCTTCAGGCACGCGCATCCACACIGCTCAGGGCATTTGCTGCGGCACGCAAGAAGCA 480  
 A P S G T A I H T A Q G I A A A R K E A  
 481 GGCATGGACGCACAGCCAGATGCGACCGAGCAGGCACCTGAGGGTTCCCGTGGCGCAAGC 540  
 G M D A Q P D A T E Q A L E G S R G A S  
 541 GTAGATGGAATCCCAGTTCaGCGAGTCCGCATGTCCGGCATGGTTGCTCAGCAGCAAGTT 600  
 V D G I P V H A V R M S G M V A H E Q V  
 601 ATCTTTGGCACCAGGGTCAGACCTTGACCATCAAGCAGGACTCCTATGATCGCACTCA 660  
 I F G T Q G Q T L T I K Q D S Y D R N S  
 661 TTTCACCAAGTGCTCTTGGTGGGTGTGCGCAACATTGCACAGCACCAGGCGCTAGTCGTA 720  
 F A P G V L V G V R N I A Q H P G L V V  
 721 GGACTTGAGCATTACCTAGGCTGTAA 747  
 G L E H Y L G L +

Nucleotide, sequence of ddh (SEQ ID NO: 1)

1 ATGCATTTCG GTAAGCTCGA CCAGGACAGT GCCACCACAA TTTTGGAGGA  
 51 TTACAAGAAC ATGACCAACA TCCGCGTAGC TATCGTGGC TACGGAAACC  
 101 TGGGACGCAG CGTCGAAAAG CTTATTGCCA AGCAGCCCGA CATGGACCTT  
 151 GTAGGAATCT TCTCGCGCCG GGCCACCCCT GACACAAAGA CGCCAGTCTT  
 201 TGATGTCGCC GACGTGGACA AGCACGCCGA CGACGTGGAC GTGCTGTTC  
 251 TGTGCATGGG CTCGCGCACC GACATCCCTG AGCAGGCACC AAAGTTCGCG  
 301 CAGTTCGCCT GCACCGTAGA CACCTACGAC AACACCCGCG ACATCCACG  
 351 CCACCGCCAG GTCATGAACG AAGCCGCCAC CGCAGCCGCG AACGTTGCAC  
 401 TGGTCTCTAC CGGCTGGGAT CCAGGAATGT TCTCATCAA CGCGTCTAC  
 451 GCAGGGSCAG TCTTAGCCGA GCACAGCAG CACACCTTCT GGGGCCCAGG  
 501 TTTGTCACAG GGCACCTCCG ATGCTTTGCG ACGCATCCCT GGGGTTCAA  
 551 AGGCCTGCGA GTACACCTCT CCATCCGAAG AAGGCCCTGA AAAGGCCGCG  
 601 CGTGGCGAAG CCGGCGACCT CACCCGAAG CAAACCCACA AGCCCAATG  
 651 CTTGCTGGTT GCGAGCGCGG CCGAGCACGA GCGCATCGAA AACGACATCC  
 701 GCACCATGCC TGATTACTTC GTTGGCTACG AAGTCGAAT CAACTTCATC  
 751 GACGAAGCAA CCTTgACgC CGAGCACACC GGCATGCCAC ACGGCGGACA  
 801 CGTGATcACC ACCGCGCACA CCGGTGGCTT CAACCAACCC GTGGAATACA  
 851 TCCTgAAGCT GGACCGAAAC CCAGATTTC CCGCTTCCTC ACAGATCGCT  
 901 TTGGGcCGCG CAGCTCACCG CATGAAGCAG CAGGGCCAAA GCGGTCCTT  
 951 CACCGTCCTC GAAGTTGCTC CATACCTGCT CTCGCCgGAG AACTTGGAtG  
 1001 ATCTGATCGC ACGGACGCTC TAA

Fig 10

097244.1.12300

SEQ ID 10: amino acid sequence of ddh.

SEQ ID

NO: 10

```

ATGATTGAGTACGCTCGACCGAGAGTGGCAGCCGATTTTGGAGATACAGAAC
100
H R F G Y L D Q D S A T T I L E D Y K N
ATGACGACATCTGGTACCTATCTGAGCTACGGAAAGCTGGGACGAGCTGCAAAAG
120
H T N I F V A I V G Y C H I C R S V E K
CTTATTGCCAAGCGCGCCACATGACCTTGTAGGATCTTCTCGCGCGCGCCACCGTC
140
L I A K Q P D H D L V G I F S R R A T L
GAGCAAGACGCCAGCTCTTGTATGTGCGCGAGCTGGACAAGACCGCGCGACGCTGAGC
160
D T K T F V F D V A D V D K H A D D V D
GTGCTGTTCTGTGATGGCTCGCGCACGACATCGCTGAGCAGCGACCAAGTTGCGG
180
V L F L C H G S A T D I P E Q A P K F A
CAGTTGCGCTGCGACCTAGACACTACGACAGCAGCGCGACATCCGACGCGCGCAG
200
Q T A C T V D T Y D N H R D I P R I R Q
GTGATGAAGCAAGCGCGACCGACGCGCGCAAGCTTGCACTGGTCTCTACCGCTGGGAT
220
V M N E A A T A A C N V A L V S T G H D
CGAGAAATGTTCTCATCAGCGCGCTTAGCGAGCGCGAGCTTAGCGGACGACGAG
240
P G M F S I N R V Y A A A V L A E R O Q
CACAGCTTCTGGGCGCAGCTTGTGACAGGCGCGACTCGGATGCTTTCGACGATCGCT
260
H T F H G P G L S Q G H S Q A L R R T P
GGCTTCAGAGGCGCTCGAGTAGCCCTCCCATCGGAAGCGCTGCAAGCGCGCGC
280
C T Q K A V Q Y T L P S E E A L E N A R
CGTCTCGAAGCGCGCGAGCTGACGGGAAGCAAAACCGACAAGCGCGCAATGTTCTGCTGGTT
300
R G E A G D L T G K Q T H K R O C T V V
GCCGAGCGCGCGGACACGAGCGCGATCGAAACGACATCGCGACGATCGCTGATTACTTC
320
A D A D N H E R I E N D I R T H F D Y F
GTTGGCTAGGAGTGAAGTCAACTTATCGAGCAAGCAAGCTTGGAGCTGAGGACAGC
340
V G Y E V E V N F I D E A T L D A E N T
GGCATGCCACGCGGGAAGAGTATCACCACCGCGGACAGCGCTGCGCTTACGACAGCC
360
G M P H G C H V I T T C D T G C G F N H T
GTGGATACATCTTGAAGCTGGAGCGGAAGCAAGATTACCGCGCTTCTACAGATCGCT
380
V E V I L K L D R N P D E T A S S Q I A
TTGCGCGCGAGCTGCGCGATGACGAGCAGCGCGCAAGCGCGCTTACCGCTGCTC

```

Fig 11A

0072244.1.12800



SEQ ID NO:10  
(cont)

```
901 ----- 960
      F G R A A H R M K G G G S G A F T V L
      GAAGTTGCTCCATACCTGCTCTCCCCGAGAACTTGGAGGATCTGATCGCAGTGGAGGTC
961 ----- 1020
      E V A P Y L L S P E N L D D L I A R D V
      TAA
1021 --- 1023
```

Fig. 11B

09722441.112800

Sequence of full length *lysA* from NRRL B-11474 (Seq ID NO:  
11); Underlined region: the priming site for *lysA* primer

ATGGCTACAGTTGAAATTTCAATGAACTTCCCGCACAGTATGCGCCACGCAATGCAGTG  
CGCCAAAGAACGGCGTTGTTCACCGTCCGTGTGTGTGCTCTCTGCTGAGCTTCGCTGAAGAA  
TAGGGAACCCUACTGTGTGTATGCTGACGAGGACGATTCCTTCTCCCGCTGTGCGGACATG  
GCTACCGCATTCGCTTGACACAGGCAATGTGCACTACGACATCCAAATGCGTTCTGACCGAG  
ACCATTGACAGTTGGTTGATGAAGAGGGGCTGGACCTGGACATTCGCTTCATCAATGAA  
CTGGGCAATTGCGCTGGCTGGTGTGTTTCGCGGCGAGCGGTATCACCGCGCACGGCAACACAC  
AAAGCGGTAGAGTTCTCTGCGCGCGTTGGTTCAAAACGGTGTGCGGCAATGTGCTGCTGAGC  
TCCGCGAGGAATTTGAACTGCTGATTAATCTCCCGCTGTGTADGACGAGATCCAGAGC  
GTGTTGATCCCGGTGAAGCCAGGTATCAGAGCCGACATCCAGATTCATCGCCACTAGC  
CACGAGACCCAGAAATTCGGAATCTCCCTTGGCATCCGCTTCGCGATTGTGAAGCAGCGAAA  
GCAGCCAACAATGCAGAGAACTTGAACCTGTGTTGGTCTGCACTGCCATGTTGGTTCCAG  
GTGTTGAGCGCGAAGGCTTCAAGCTGGCAGCAGAGCGCGTGTGGCGCTGTACTCAGAC  
ATCCACAGCGGAATAGTGTGCGCTTCTGTAGCTGGACCTCGGTGGCGAATAGGACATC  
GCTTACATCTGAGATGAGGAACCATCAACCGTCCAGAGATGCGCTCCGACTGCTACCC  
GCATCCGGAATAATGGCAGCGGAATAGGCTATGACGACCAACCTGCTGTGTGAGCCG  
GGCGCGCTATCCGAGCGCCCTCCACCGTGACCTCTAGAGAGTGGGACCAACAAAAC  
GTCCACGTAGACGACGACAAAACCGCGCGTACCTAGCGCGCTGACGAGGCGATGTCCGAC  
AAACTCCGCGCAGCACTCTACCGCTCCGAAATACGACGCGCGCGGTAGTATCCCGCTTCCG  
GAGGAGACCCGATAGCACCCGATGTGTGGCTCCGACTGCGAATCCGCGATATCCCTG  
ATCAACGATGAATCTACCGATCTGATGATACCGAGCGGAGCTCTGCTGCTGACGCG  
ACCGGCGCATACTGCTACGCGCATGAGCTCCCGCTACAAACGCTTCAACAGCGCGCGCCCTC  
GTGTCGCTCCCGCTGGCAGCTCCCGCTCATGCTGCGCGCGGAACCGCTGACGACATC  
CTCTCACTAGAGGCTAA

Fig. 12

0972441.112800

Full length sequence of LysA (NRRL-B11474)  
 DIAMINOPIMELATE DECARBOXYLASE (Lys A) SEQ ID NO:12

MATVENFNELPANVWERNVVRQEDGVTVAGVPLPDLAEEYGTPLFVDEDDFRSRCDM  
 ATAFGGEGNVHYASKAFLETKTARWVDEBGLALDIASINELGIALAAGFPASRTAHGNN  
 KGVFELRALVQNGVGHVVLDSAQEELLDYVAAGEGKIQDVLIRVKPGLEARTHEFLATS  
 HEDQKFGFSLASGSAFEAAKAANNAENLNVLNGLHCHVGSQVFDARGFKLAERVLGLYSQ  
 IHSELGVALFELDLGGGTYGIAYTADEEPLAVAEVASEDLLTAUGKMAELGIDAPTVLVFP  
 GRALAGFSTVTIYEVGGTYKNVHVDODKTRAYVAVDGKMSDNIRLPALYGYEYDARVVSREA  
 EGDVSTRIVGSECESSGDILINDEITYPSDITSGDFLAALATGAYCYAMSSRYNAFTRPAV  
 VSVRAGSSRLMLRRETLDLILSLA

Fig. 13

0072441.12800

1 ATGGCTACAG TTGAAAATTT CAATGACTT CCGGCACAG TATGCCACG  
 51 CAATGCCGTG CGCCAGAGG ACGGCTTTGT CACCGTCGCT GGTGTGCTC  
 101 TGCCTGACCT CGCTAAGAA TACGGACCC CACTGTTGCT AGTCGACGAG  
 151 GACGATTTCG GTTCCCGCTG TCGCGCATG GCTACCGCAT TCGGTGGACC  
 201 AGGCAATGTG CACTACGCAT CTAAGCGTT CCGACCAAG ACCATTGCAC  
 251 GTTGGGTTGA TGAAGAGGGG CTGGCACTGG ACATTGCATC CATCAACGAA  
 301 CTGGGCATTG CCTGGCCGC TGGTTCCGC GCCAGCGTA TCACCGCGCA  
 351 CGCAACAAC AAAGCGTAG AGTTCTGCG CGCGTTGTT CAACACGGTG  
 401 TGGGACAGT GGTGCTGGAC TCCGCACAG AACTAGAACT GTTGGATTAC  
 451 GTTGCCGCTG GTGAAGGCA GATTCAAGAC GTGTTGATCC GCGTAAGCC  
 501 AGGCATCGAA GCACACACC ACGAGTTCAT CGCCACTAGC CACGAGACC  
 551 AGAAGTTCGG ATTCTCCCTG GCATCCGTT CGCATTCGA AGCAGCAAAA  
 601 GCGCCAAACA ACGCAGAAAA CCTGACCTG GTTGGCTGC ACTGCCACGT  
 651 TGGTCCAG GTGTTGACG CCGAGGCTT CAAGCTGGCA GCAGAACGG  
 701 TGTGGGCTT GTACTCAG ATCCAGCGC AACTGGCGT TGCCCTTCCT  
 751 GAATGGATC TCGGTGGCG ATACGCTATT GCCTATACCG CAGCTGAAGA  
 801 ACCACTCACC GTGCAGAG TTGCTCCGA CCGTCTACC GCAGTCGGAA  
 851 AATGGCAGC GGAATAGG ATCGACGAC CAACCGTGT TGTGAGCC  
 901 GCGCGGCTA TCGCAGGCC CTCACCGTG ACCATCTAG AAGTCGGCAC  
 951 CACCAAGAC GTCCAGTAG ACGACGAA AACCCGCGT TACATCGCG  
 1001 TGGACGGAG CATGTCCGAC AACATCGCC CAGCACTTA CGGCTCCGAA  
 1051 TACGACGCC GCGTAGTAT CCGCTCGCC GAAGGAGAC CAGTAAGCAC  
 1101 CCGCATCGT GGTCCCACT CGGAATCGG CGATATCCTG ATCAACGATG  
 1151 AATCTACCC ATCTGACAT ACCAGTGGG ACTTCCTTG ACTCGCAGCC  
 1201 ACCGGCGCAT ACTGCTACG CATGAGTCC CGCTACAACG CCTTCACAG  
 1251 GCCCGCGTC GTGTCCGTC GCGCTGGC CTCGCGCTC ATGCTGCGCC  
 1301 CGCAACGCT CGACGACATC CTCCTACTAG AGGCATAA

Fig. 14

0972441 112800

Fig 15A

097441-1300

Fig 15B

Fig 15C RD NO: 14 Lys A (pRS6)

(cont.)

CTCTCACTAGAGGCATAA  
1321 ----- 1338

L S L E A \*

Fig 15C

09722441.12800

Nucleotide. sequence of orf2 in dapBA operon SEQ ID NO: 15

1 GTGGCCGAAC AAGTTAAATT GAGCGTGGAG TTGATAGCGT GCAGTTCCTT  
51 TACTCCACCC GCTGATGTTG AGTGGTCAAC TGATGTTGAG GCGCGGAAG  
101 CACTCGTCGA GTTGGGSGT CGTGCTGCTT ACGAATCTTT TGATAGACCG  
151 AACCCTCGAA CTGCTTCCAA TGCTGCGTAT CTGGCCGACA TCATGGAAGT  
201 GGGGCACACT GCTTTGCTTG AGCATGCCAA TGCCACGATG TATATCCGAG  
251 GCATTTCTCG GTCCGCGACC CATGAATTGG TCCGACACCG CCATTTTCTC  
301 TTCTCTCAAC TGTCTCAGCG TTTCTGTCAC AGCGAGAGAT CGGAAGTAGT  
351 GGTGCCCACT CTCATCGATG AAGATCCGCA GTTGCCTGAA CTTTTCATGC  
401 ACGCCATGGA TGAGTCTCGG TTCGCTTTCA ATGAGCTGCT TAATGCGCTG  
451 GAAGAAAAAC TTGCCGATGA ACCGAATGCA CTTTARGGA AAAAGCAGGC  
501 TCGTCAAGCA GCTCGCGCTG TGCIGCCCAA CGCTACAGAG TCCAGAATCG  
551 TGGTGTCTGG AAACCTCCGC ACCTGGAGGC ATTTTATTGG CATGCCAGCC  
601 AGTGAACRTG CAGACGTGCA RATCGCGGAA GTAGCGGTAG GATGTTTAG  
651 AAAGCTGCAG GTAGCAGCGC CAACTGTTTT CGGTGATTTT GAGATTGAAG  
701 CTTTGGCAGA CGGATCGCAA ATGGCAACAA GCCCGTATGT CATGGACTTT  
751 TAA

Fig 16



SEQ ID No: 16

ORF2 amino acid sequence

1 GTGGCGGAACAAGTTAAATTGAGCGTGGAGTTGATAGCGTGCASTTCTTTTACTCCACCC 60  
 M A E Q V R L S V E L I A C S S F T F P  
 61 GCTGATGTTGAGTGGTCAACTGATGTTGAGGGCGCGGAGCACTCGTCGAGTTTGGGGT 120  
 A D V E W S T D V E G A E A L V E F A G  
 121 CGTGCTGCTACGAACTTTTGATAAGCCGAACCTCGAACTGCTTCCAATGCTCGGTAT 180  
 R A C Y E T F D K P N P R T A S N A A Y  
 181 CTGGCCACATCATGGAAGTGGGGCACACTGCTTTGCTTGAGCATGCCAATGCCACGATG 240  
 L R H I M E V G H T A L L E H A N A T M  
 241 TATATCCGAGGCATTTCTCGGTCCGCGACCCATGAATGGTCCGACACCGCCATTTTCC 300  
 Y I R G I S R S A T H E L V R H R H F S  
 301 TTCTCTCAACTGTCTCAGCGTTTCGTGCACACGGGAGAATCGGAAGTAGTGCTGCCCACT 360  
 F S Q L S Q R F V H S G E S E V V V P T  
 361 CTCATCGATGAAGATCCGCGAGTTGCGTGAACTTTTCATGACGCCCATGGATGAGTCTCGG 420  
 L I D E D P Q L R E L F M H A M D E S R  
 421 TTGGCTTTCAATGAGCTGCTTAATGCGCTGGAAGAAAACTTGGCGATGAACCGAATGCA 480  
 F A F N E L L N A L E E K L G D E P N A  
 481 CTTTTAAGGAAAAAGCAGGCTCGTCAAGCAGCTCGCGGTGTGCTGCCCAACGCTACAGAG 540  
 L L R K K Q A R Q A A R A V L F N A T E  
 541 TCCAGAATCGTGGTCTCTCGAATCTTCCGACCTGGAGGCATTTCATTGGCATCGGAGCC 600  
 S R I V V S G N F R T W R H F I G M R A  
 601 AGTGAACATGCAGACGTCGAATCCGCGAAGTACCGGTAGGATGTTAAGAAAGCTGCAG 660  
 S E H A D V E I R E V A V G C L R K L Q  
 661 GTAGACGCGCAACTGTTTTCGGTGATTTGAGATTGAACTTTGGCAGACGGATCGCAA 720  
 V A A P T V F G D F E I E T L A D G S Q  
 721 ATGGCAACAAGCCGATGTCATGGACITTTAA 753  
 M A T S P Y V M D F \*

Fig 17

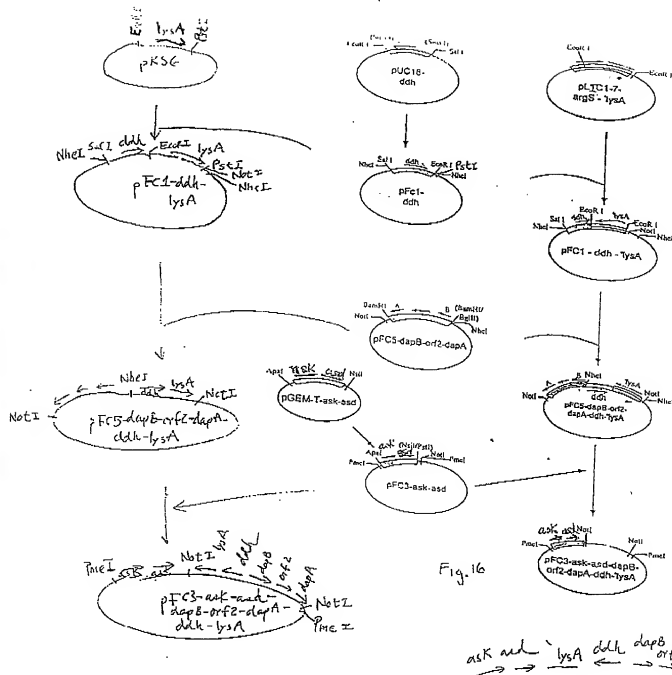


Fig. 16

Fig. 18

ask ask *lysA* *dch* *dapB* *orf2* *dapA*

ATCC 13032	1	50
N13		
ATCC 21529		
Consensus	HALVVQKYGG SSLESAERIR HVAERIVATK KAGND/VVVC SANGOTTDEL	
ATCC 13032	51	100
N13		
ATCC 21529		
Consensus	LELAADVNPV PPAREMDMLL TAGERISNAL VMAIESLGA EAQSTGSGQA	
ATCC 13032	101	150
N13		
ATCC 21529		
Consensus	GVLTTERRHGN ARIVDVTPGR VREALDEGKI CIVAGFGVN KETRDVTTLG	
ATCC 13032	151	200
N13		
ATCC 21529		
Consensus	RGGSDDTTAVA LAAALNADVC EIYSDVDGVY TADPRIVPNA QZLEKLSFEE	
ATCC 13032	201	250
N13		
ATCC 21529		
Consensus	MLELAAVGSK ILVLRVVEYA RAFNVPLRVH SSYSNDPOTL IAGSMEDIPV	
ATCC 13032	251	300
N13		
ATCC 21529		
Consensus	EEAVLTGVAT DKSEAKVTVL GISDKPGEIA KVFRALADAE IMIDMVLQNV	
ATCC 13032	301	350
N13	S	G
ATCC 21529	A	D
Consensus	SSVEDGTTDI TPTCPRADGR RAKELKRLQ VGNWNTNVLV DQVGVSLV	
ATCC 13032	351	400
N13	T	
ATCC 21529	T	
Consensus	GACMKSHPGV TAEFMEALRD VVWIELIST SEIRISVLIR EDDLDAARA	
ATCC 13032	401	421
N13		
ATCC 21529		
Consensus	LHEQFQLGGE DEAVVIAGTS R	

Fig. 19

09722441.11360

HpaI - PvuII fragment comprising the P1 promoter

AACCGGTGTGGAGCCGAACCATTCGCGGAGGCTGCACTGCAACGAGGTGCTAGTTTGGTACATGGCTTCTG  
GCCAGTTCATGGATTGGCTGCCGAAGAAGCTATAGGCATGCCACCGAGGGCCACCGGAGTTACCGAAGAT  
GGTGCCGTGCTTTTCGCCITGGGCAAGGACCTTGACAAAGCCACCGCTGATATCGCCAAAGTGAGGGATCAG  
AATAGTGCATGGGCACGTCGATGCTGCCACATTGACGCGAGGCAATATCTACCTGAGGTGGGCATTCTCC  
CAGCGGATGTTTCTTGCGCTGCTGCACTGGGCATTGATACCAAAAGGGGCTAAGCGCAGTTCGAGGCGG  
CAAGAAGTCTACTACCTTTTTTATTGTGCAACGGGGCATTACGGCTCCAAGGACGTTTGTTTCTGGGTCA  
GTTACCCCAAAAAGCATATACAGAGACCAATGATTTTTCATTAAAAAGGCAGGGATTTGTATATAAGTATGG  
GTGCTATTCTGTGCGACGGGTGTACCTCGGCTAGAATTTCTCCCATGACACAG

Fig. 20 (SEQ ID NO: 17)

# Making pFCI-dhh-P<sub>lysA</sub>

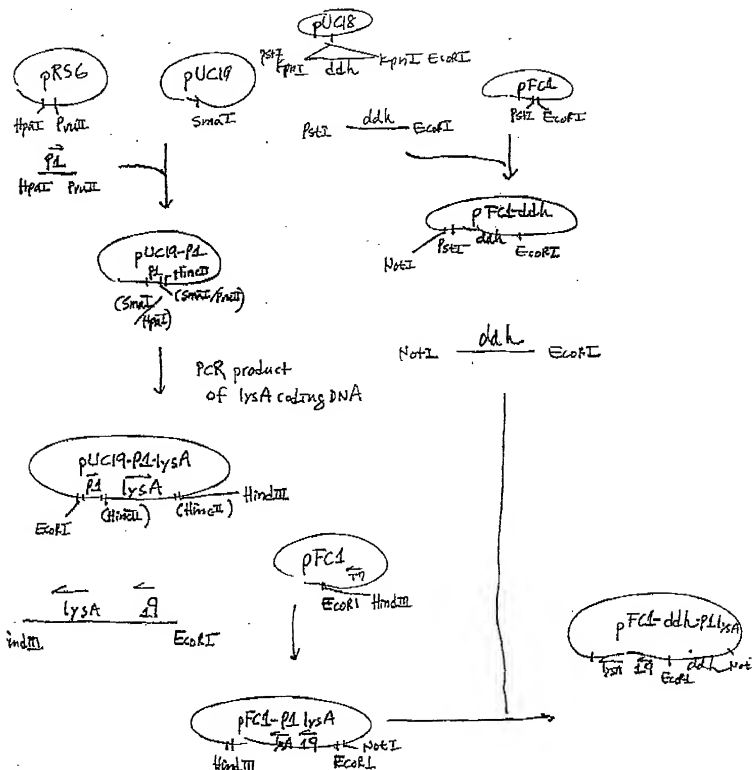


Fig. 21A

# Making pDEla2-KDABHPL

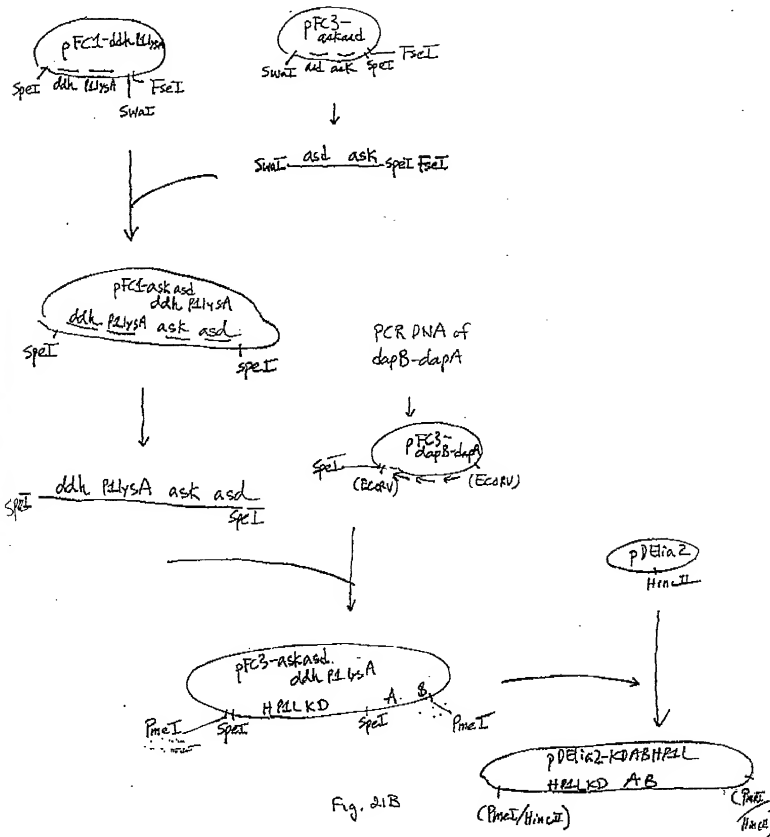
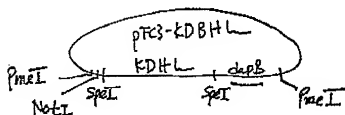
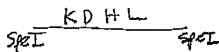
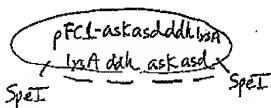
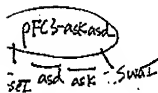
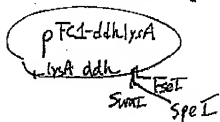


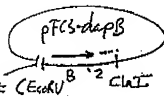
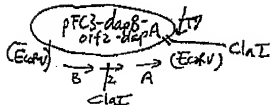
Fig. 21B

# Making pDEla2FCS-KDBHL

(from Fig 16)



PCR DNA  
of dapB-dapA



(the truncated  
orf 2 ----)

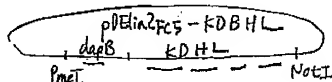
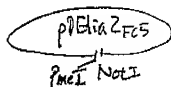
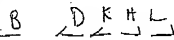


Fig. 22

orientation:



nucleotide sequence of truncated ORF2

SPR ID NO: 18

```

1  GTGGCCGAAC  AAGTTAAATT  GAGCGTGGAG  TTGATAGCGT  GCAGTTCTTT
51  TACTCCAGCC  GCTGATGTTG  ACTGGTCAAC  TGATGTTGAG  GCGCGGGAAG
101  CACTCGTGGG  GTTGGCGGGT  CGTGCTGCT  ACGAACTTT  TGATAAGCCG
151  AAGCCTCGAA  CTGGTCCAA  TGTGCGTAT  CTGGGCCACA  TCATGGAAGT
201  GGGGCACACT  GCTTGCTTG  AGCATGCCAA  TCGCACGATG  TATATCCGAG
251  GCATTTCTCG  GTCCGCGGCC  CATGAATTGG  TCCGACACCG  CCATTTTCC
301  TTCCTCAAC  TCTCTCAGCG  TTTGCTGCAC  AGCGAGAAAT  CGGAAGTAGT
351  GTGCCCACCT  CTCAT

```

Fig. 23

09722441.112800



Truncated ORF-2

amino acid sequence

GTGGCCGAACAAGTTAARTTGAAGCGTGAAGTTGATAGCGTGCAGTTCTTTTACTCCACCC  
1 ----- 60  
M A E Q V K L S V E L I A C S S F T P  
CGTGATGTTGACTGCTCAACTGATGTTGAGCGCGGTAAGCACTGTCGAGTTTGGGGT  
61 ----- 120  
A D V E W S T D V E G A E A L V E F A G  
CGTGCCTGCTACGAACATTTTGATAAGCGGAACCTCGAAGTCTTCCAATGCTGCGATAT  
121 ----- 180  
R A C Y E T F D K P N P R T A S N A Y  
CTGCGCCACATCATGGAAGTGGGGCACACTGCTTTGCTGAGCATGGCCATGCCACGATG  
181 ----- 240  
L R H I M E V G H T A L L E H A N A T M  
TATATCGAGGCAITTTCTCGTCCGCGACCCATGAATGGTCCGACACGCCATTTTTC  
241 ----- 300  
Y I R G I S R S A T H E L V R H R H F S  
TTCTCTCAACTGTCTCAGCGTTTCGTGCACAGCGAGAAATCGGAAGTAGTGGTGCCCACT  
301 ----- 360  
F S Q L S Q R F V H S G E S E V V V P T  
CTCAT ...  
361

L (I)

SEK ID NO: 20

Sequence of truncated Lys A (Lys A)

(NRRL-B11474)

ATGGCTACAGTTGAATAATTCAATGAACATCCCGCACAGGTATGGCCACGCCAATCGCATG  
 CGCCAAAGAGACGGCGTTGTACCCGTGCTGGTGTGCTCTGCTGACCTGCTGAAGAA  
 TAGCGAACCACACTGTTGTAGTCGAGAGGAGCAGATTTCCTTCCTCCCTGCGCGACATG  
 GCTACCGCATTCGGTTCGACACGGCAATGTGCACTACGCATCCAAAGCGTTCTGTACCAAG  
 ACCATTGCACGTGGGTGTATGAAGAGGGGCTGCGACTGACATTCGCGTCCATCAATGAA  
 CTGGGCAATGCCCTGGCCGCTGGTTCCCGCCAGCGGTATCACCCGACCGCAACAAAC  
 AAAGGCGTAGAGTTCCTGCGCGCTTGGTTCAAAACGGTGTGCGGCATGTGGTCTGGAC  
 TCCGCGCAGGAATCGAAGCTGCTGGATTACGTTGCCCGTGGTGAAGCAAGATCCAGGAC  
 GTGTTGATCCCGGTGAAGCCAGGTATTCGAAGCCCAACCCACGATTCATCGCCACTAGC  
 CACGAAGCCAGAAATTCGGATTCTCCCTGGCATCCGGTTCCCGCATTCGAAGCAGCGAAA  
 GCAGCCCAACATGCAGAGAACTTGAACCTGGTTGGTCTGCACTGCCATGTGGTTCCAG  
 GTGTTGAGACGCCGAGGCTTCAAGCTGACGACAGAGCCGGTTTGGCCCTGTACTACAG  
 ATCCACAGCGAACTAGGTGTGCGCCCTTCTGAGCTGGACCTGGTGGCGGATACGGGCATC  
 GCCTACACTGCAGATGAGGAACCACTCAAAGTGGCAGAAAGTGGCCCTCCGACCT

Fig. 25

Truncated sequence of *lysA* (NRRL-B11474).

DIAMINOPIMELATE DECARBOXYLASE (*lysA*) SEQ ID NO: 21

MATVENFNELPAHVWERNAVRQEDGVTVAGVPLDLASEYGTPLFVVDDEDFSRCRDM  
ATAFGGGEGNVHYASKAFLTKTLARVVDDEGLADIASINELGIALAAGFPASRITAHGNN  
KGVEFLRALVQNGVGHVVLDSAQSTLELLDYVAAGEGKIQDVLIRVKPGIEATHEFIATS  
BEDQKFGFSLASGSFAFAAKAANNANLNVLGHCIVGSQVFDAGCFKLAERNVLGLYSQ  
IHSELGVALPELDLGGGYGIATADDEPLNVAEVASDIL

Fig. 26

0372441.1.2800